

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:30 ; Search time 70 Seconds
(without alignments)
5011.310 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSLGVLSPVLLQLAT.....PAYPTESCHLSLVAFVPCVL 907

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1980s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4702	100.0	907	2	AAW93889 Human HG3
2	4702	100.0	907	3	AAY90682 Human G P
3	4702	100.0	907	6	ABP81968 Human G P
4	4702	100.0	907	6	ABO06467 Human G-P
5	4702	100.0	907	7	ADC22783 Human G P
6	4702	100.0	907	7	ADSE9150 Human Pro
7	4702	100.0	907	7	ADSE9153 Human Pro
8	4702	100.0	907	7	ADG42628 Human G P
9	4702	100.0	907	7	ADH14256 Human HG3
10	4702	100.0	907	7	ADN40013 Cancer/an
11	4702	100.0	907	7	ADN39531 Cancer/an
12	4702	100.0	907	7	ADN39628 Cancer/an
13	4702	100.0	907	8	ADO29408 Human GPC
14	4702	100.0	907	8	ADQ80369 G protein
15	4702	100.0	907	8	ADR67868 Human HG3
16	4701	100.0	1145	7	ADP70480 Orphan re
17	4696	99.9	907	3	AAY90687 Human mut
18	4696	99.9	907	7	ADC22797 Human G P
19	4696	99.9	907	7	ADH14270 Mutated h
20	4691	99.8	907	7	ADG42629 Human G P
21	4556	96.9	883	7	ADB80464 Ovarian c
22	4556	96.9	883	7	ADN40012 Cancer/an
23	4556	96.9	883	7	ADN39166 Cancer/an
24	4048	86.1	907	8	ADG75449 Mouse orp
25	4048	86.1	907	8	ADO29409 Mouse GPC

26	3598	76.5	693	2	AAAY42169	Aay42169 Human LGR
27	3564	75.8	692	2	AAW93890	Aaw93890 Human HG3
28	2534	53.9	967	5	AAG66138	Aag66138 Mouse LGR
29	2534	53.9	967	7	ADN02240	Adn02240 Mouse lar
30	2534	53.9	967	8	ADK19405	Adk19405 Mouse LGR
31	2527.5	53.8	967	7	ADC16693	Adc16693 Human G-P
32	2527.5	53.8	1205	7	ADF70441	Adf70441 Orphan re
33	2525.5	53.7	967	5	AAG66141	Aag66141 Human LGR
34	2525.5	53.7	967	7	ADC16691	Adc16691 Human G-P
35	2525.5	53.7	967	7	ADD89091	Add89091 TAT289.1
36	2525.5	53.7	967	7	ADN02249	Adn02249 Human ful
37	2525.5	53.7	967	8	ADK19414	Adk19414 Human LGR
38	2525.5	53.7	967	8	ADQ15074	Adq15074 Human can
39	2525.5	53.7	1005	6	ABJ37069	Abj37069 Human bre
40	2494	53.0	1014	7	ADC16699	Adc16699 Human G-P
41	2492.5	53.0	928	7	ADC16711	Adc16711 Human G-P
42	2492.5	53.0	928	7	ADG42625	Adg42625 Human VTS
43	2492	53.0	1014	7	ADC16697	Adc16697 Human G-P
44	2465	52.4	940	5	Aau79941	Aau79941 Human orp
45	2450.5	52.1	940	7	ADG42573	Adg42573 Novel hum

ALIGNMENTS

RESULT 1

AAW93889
ID AAW93889 standard; protein; 907 AA.

XX AC AAW93889;

XX DT 25-JUN-1999 (first entry)

XX DE Human HG38 protein.

XX KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain;

XX KW endocrine system; skeletal muscle; spinal cord; placenta; development;

XX KW receptor activity modulator.

XX OS Homo sapiens.

XX PN WO9915660-A1.

XX PD 01-APR-1999.

XX PF 24-SEP-1998; 98WO-US019979.

XX PR 24-SEP-1997; 97US-0059863P.

XX (MERI) MERCK & CO INC.

XX Liu Q, Bailey WJ, McDonald TP;

XX WPI; 1999-254711/21.

XX N-PSDB; AAX23980.

XX Human G-protein coupled glycoprotein hormone receptor HG38.

XX Claim 1a; Fig 2; 74pp; English.

XX This invention describes a novel human G-protein coupled glycoprotein hormone receptor, HG38. Glycoprotein hormone receptors are important in the endocrine system and HG38 may be involved in development and function of the skeletal muscle, spinal cord, placenta and to a lesser extent, the brain. The transgenic animal may be useful for studying tissue and temporal specific expression or activity of the HG38 receptor, as well as for studying the ability of a variety of compounds to act as modulators of HG38 receptor activity

SQ Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 2; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDTSLGVLISLPVLLQLATGSSPRSGVLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db	1 MDTSLGVLISLPVLLQLATGSSPRSGVLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Qy	61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRLEELRAGNALTYPKGAFTGLSKLVLM 120
Db	61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRLEELRAGNALTYPKGAFTGLSKLVLM 120
Qy	121 LQNNQLRHPVTEALQNLRSLSRLDANHSISYVPPSCFSGLSRLHRLWLDNALTETIPVQ 180
Db	121 LQNNQLRHPVTEALQNLRSLSRLDANHSISYVPPSCFSGLSRLHRLWLDNALTETIPVQ 180
Qy	181 AFRSLSALOQMTALNKHHTPDYAFGNLSLVVLHNNRHSIGKKCPGLHSLETLTD 240
Db	181 AFRSLSALOQMTALNKHHTPDYAFGNLSLVVLHNNRHSIGKKCPGLHSLETLTD 240
Qy	241 LNYNNLDEFPPTAIRTLNKLGLGFSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
Db	241 LNYNNLDEFPPTAIRTLNKLGLGFSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
Qy	301 FOHLPELRTLTLNGASQITTEPDDLTGTANLESFLTGAQISSLPQTVNCQNPNIQVLDLS 360
Db	301 FOHLPELRTLTLNGASQITTEPDDLTGTANLESFLTGAQISSLPQTVNCQNPNIQVLDLS 360
Qy	361 YNLLEDLPSFSVCOKLQKIDLRHNEIYEIKVDTFQQLLSLSRLANWKIAIHPNAPST 420
Db	361 YNLLEDLPSFSVCOKLQKIDLRHNEIYEIKVDTFQQLLSLSRLANWKIAIHPNAPST 420
Qy	421 LPSLIKLDLSNLLSSFPITGLHGLTHLKTGNHALQSLISSENFPELPKVIEMPVAYQCC 480
Db	421 LPSLIKLDLSNLLSSFPITGLHGLTHLKTGNHALQSLISSENFPELPKVIEMPVAYQCC 480
Qy	481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Db	481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Qy	541 CSPSPGPKPCHEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db	541 CSPSPGPKPCHEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy	601 AVNMLTGVSAAVLGAVDAFTGSPARHAGWENGVCCHVIGFLSIFASESSVFLTLAAL 660
Db	601 AVNMLTGVSAAVLGAVDAFTGSPARHAGWENGVCCHVIGFLSIFASESSVFLTLAAL 660
Qy	661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKYGASPICLPLPFGEP 720
Db	661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKYGASPICLPLPFGEP 720
Qy	721 STMGVYVALIILNSICFLMMTIAVTKLYCNLDKGDLENIPDCSMVKHTALLFTNCILNC 780
Db	721 STMGVYVALIILNSICFLMMTIAVTKLYCNLDKGDLENIPDCSMVKHTALLFTNCILNC 780
Qy	781 PVAFLSFLSINLTFISPEVIFKILLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTYV 840
Db	781 PVAFLSFLSINLTFISPEVIFKILLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTYV 840
Qy	841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSTITVDLPSSVPSPAYPVPTESCHLSS 900
Db	841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSTITVDLPSSVPSPAYPVPTESCHLSS 900
Qy	901 VAFVPCL 907
Db	901 VAFVPCL 907

RESULT 2
AA90682
XX AA90682 standard; protein; 907 AA.
AC
XX AA90682;

DT 21-AUG-2000 (first entry)
XX Human G protein-coupled receptor HG38.
XX
XX G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening; agonist;
KW antagonist.
XX Homo sapiens.
XX WO200022129-A1.
PD 20-APR-2000.
XX
XX 12-OCT-1999; 99WO-US023938.
XX 13-OCT-1998; 98US-00170496.
PR (AREN-) ARENA PHARM INC.
XX Behan DP, Chalmers DT, Liaw CW;
XX WPI; 2000-329165/28.
DR N-ESDB; AAA30770.
XX Non-endogenous constitutively activated human G protein-coupled
PT receptors, useful for identifying agonists for use as pharmaceutical
PT agents.
XX
PS Example 1; Page 317-320; 341pp; English.
XX The invention relates to constitutively active, non-endogenous versions
CC of endogenous human orphan G protein-coupled receptors (GPCRs, AA90643-
CC AA90677 and AA90683-Y90687), and to DNA encoding them (AA90709-A30743
CC and AA90775-A30779). The mutant proteins of the invention contain a
CC mutation in a portion of the protein comprising intracellular loop 3
CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,
CC is substituted for an endogenous residue in IC3 at a position 16 amino
CC acids N-terminal of an endogenous proline in TM6 to form a sequence X-
CC (AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg or
CC Ala, and is preferably Lys. When the endogenous residue at this position
CC is Lys, this residue is replaced by His, Arg or preferably Ala. The 15
CC amino acid stretch between the substituted amino acid and the Pro may be
CC endogenous, non-endogenous, or a mixture of endogenous and non-endogenous
CC residues. The constitutively active GPCRs are useful for identifying
CC antagonists, agonists and partial agonists for use as pharmaceutical
CC agents. The mutant proteins are also useful in research settings for
CC elucidating the roles of the receptors in normal and diseased conditions.
CC Antagonists for a particular GPCR are useful for treating diseases and
CC disorders associated with that receptor. Because the novel mutant GPCRs
CC are constitutively active, they can be used directly for screening of
CC compounds without the need for endogenous ligands. The present sequence
CC represents a human wild-type GPCR referred to in an exemplification of
CC the invention
XX
SQ Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 3; Length 907;	
Best Local Similarity 100.0%; Pred. No. 0;	
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDTSLGVLISLPVLLQLATGSSPRSGVLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db	1 MDTSLGVLISLPVLLQLATGSSPRSGVLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Qy	61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRLEELRAGNALTYPKGAFTGLSKLVLM 120
Db	61 PSNLSVFTSYLDLSNNNISQLLPNPLPSLRLEELRAGNALTYPKGAFTGLSKLVLM 120
Qy	121 LQNNQLRHPVTEALQNLRSLSRLDANHSISYVPPSCFSGLSRLHRLWLDNALTETIPVQ 180
Db	121 LQNNQLRHPVTEALQNLRSLSRLDANHSISYVPPSCFSGLSRLHRLWLDNALTETIPVQ 180

QY 181 AFRSLSALQAMTALNKHIIHIDYAFGNLSLVVHLHNNRTHSLGKKCFDGLHSLETLD 240
Db 181 AFRSLSALQAMTALNKHIIHIDYAFGNLSLVVHLHNNRTHSLGKKCFDGLHSLETLD 240
QY 241 LNNYNNLDEPPTAIRLTSNLKELGPHSNIRSIPEKAFVGNPSLITTHFYDNPFIQFVGRSA 300
Db 241 LNNYNNLDEPPTAIRLTSNLKELGPHSNIRSIPEKAFVGNPSLITTHFYDNPFIQFVGRSA 300
QY 301 FOHLPELRLTLNGASQITEFPDLTGATNLESILITGAQISSLPQVNCNQLNQLVLDLS 360
Db 301 FOHLPELRLTLNGASQITEFPDLTGATNLESILITGAQISSLPQVNCNQLNQLVLDLS 360
QY 361 YNLLLEDLPFSVCOKLOKIDLRHNEYIEIKVDTFOQLLSRLSNLAWNKIAIHNPAFT 420
Db 361 YNLLLEDLPFSVCOKLOKIDLRHNEYIEIKVDTFOQLLSRLSNLAWNKIAIHNPAFT 420
QY 421 LPSLIKDLSSNLLSFPITGLHGLTHLKLGTGNHALQSLSISSENPELKVIMPYAYQCC 480
Db 421 LPSLIKDLSSNLLSFPITGLHGLTHLKLGTGNHALQSLSISSENPELKVIMPYAYQCC 480
QY 481 AFGVCENAYKISQNMWKGNSMDDLHKKDAGMFOAQDERDLEFLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNMWKGNSMDDLHKKDAGMFOAQDERDLEFLDPEEDLKALHSVQ 540
QY 541 CSPSPGPPKCEHLLDGLWIRIGWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPPKCEHLLDGLWIRIGWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLADFTGSPARHGAWENGCHVIGFLSIPASSSSVFLTLAAL 660
Db 601 AVNMLTGVSASVAVLADFTGSPARHGAWENGCHVIGFLSIPASSSSVFLTLAAL 660
QY 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLATMAAVALLGSGKYGASPLCLPFGSP 720
Db 661 ERGFSVKYSAPETKAPFSSSLKVIILLCALLATMAAVALLGSGKYGASPLCLPFGSP 720
QY 721 STMGYVALILNLSICFLMWTATYKLYCNLDKGLNLDWDCSMVKHIALLFITNCILNC 780
Db 721 STMGYVALILNLSICFLMWTATYKLYCNLDKGLNLDWDCSMVKHIALLFITNCILNC 780
QY 781 PVAFISFSSLINLTFTSPRVIKFIILLVVVPLPACLNPLIYILFNPHFKEDLVSLRKQTV 840
Db 781 PVAFISFSSLINLTFTSPRVIKFIILLVVVPLPACLNPLIYILFNPHFKEDLVSLRKQTV 840
QY 841 WTRSKHPSLMSINSDVQKSCDSTQALVFTSSSITYDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVQKSCDSTQALVFTSSSITYDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPCL 907
Db 901 VAFVPCL 907

RESULT 3
ID ABP81968 standard; protein; 907 AA.
XX
AC ABP81968;
XX
DT 04-MAR-2003 (first entry)
XX
DE Human G protein-coupled receptor GPR49 protein SEQ ID NO:422.
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;

KW ulcer.
XX Homo sapiens.
XX WO200261087-A2.
XX 08-AUG-2002.
PF 19-DEC-2001; 2001WO-US050107.
XX 19-DEC-2000; 2000US-0257144P.
XX (LIFE-) LIFESPAN BIOSCIENCES INC.
XX Burner GC, Roush CL, Brown JP;
XX WPI: 2003-046718/04.
XX N-PSDB; AB242816.
PT New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.
XX Disclosure; Fig 1; 523pp; English.
CC The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (1) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial lung, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 6; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSELGVLLSLPVLQLATGSSPRSGVLLRGCPHCEPDGRMLLRVDCSDGLSEL 60
Db 1 MDTSELGVLLSLPVLQLATGSSPRSGVLLRGCPHCEPDGRMLLRVDCSDGLSEL 60
QY 61 PSNLSVFTSYLDLSNMNLSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSILKVL 120
Db 61 PSNLSVFTSYLDLSNMNLSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSILKVL 120
QY 121 LQNNQLRHVPTEALQNLRSLSLRDANHISVPPSPCSGLHSRLHMLDDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSLRDANHISVPPSPCSGLHSRLHMLDDNALTEIPVQ 180
QY 181 AFRSLSALQAMTALNKHIIHIDYAFGNLSLVVHLHNNRTHSLGKKCFDGLHSLETLD 240
Db 181 AFRSLSALQAMTALNKHIIHIDYAFGNLSLVVHLHNNRTHSLGKKCFDGLHSLETLD 240

102 e 7
Not species for

QY 241 LYNVNLDEFPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIDFVGRSA 300
DB 241 LYNVNLDEFPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIDFVGRSA 300
QY 301 FOHLPELRTLTNGASQITTEPDLTGTANLESRLTGTGAQISSLPQTVCNQLPNIQVLDLS 360
DB 301 FOHLPELRTLTNGASQITTEPDLTGTANLESRLTGTGAQISSLPQTVCNQLPNIQVLDLS 360
QY 361 YNLLEDLPSPVSVCKLOKIDLRHNEIYEIKVDTFQQLLSRLSLANWKIAIHPNAPST 420
DB 361 YNLLEDLPSPVSVCKLOKIDLRHNEIYEIKVDTFQQLLSRLSLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISENPFPELKIEMPVAYOCC 480
DB 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISENPFPELKIEMPVAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDLHKKADAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDLHKKADAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPECHLLDGLHIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPKPECHLLDGLHIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSPFARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSVAVLGADFTGSPFARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
QY 661 ERGSVKYSKAFETKAPFSSLKVAILLCALLATMAAVPLLGSSKYGASPLCLPFGEP 720
DB 661 ERGSVKYSKAFETKAPFSSLKVAILLCALLATMAAVPLLGSSKYGASPLCLPFGEP 720
QY 721 STMGYVALIILNSLCFLMWTIATKLYCNLDKGDLENWPCSMVKHIALLLFNCILNC 780
DB 721 STMGYVALIILNSLCFLMWTIATKLYCNLDKGDLENWPCSMVKHIALLLFNCILNC 780
QY 781 PVAFLSFSSNLNLTPISEVIFKILLVVPACLNPLLYILFNPHKEDLVLSRKQTYV 840
DB 781 PVAFLSFSSNLNLTPISEVIFKILLVVPACLNPLLYILFNPHKEDLVLSRKQTYV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVFTSCHLSS 900
DB 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVFTSCHLSS 900
QY 901 VAFVPCCL 907
DB 901 VAFVPCCL 907

RESULT 4
ABO06467
ID ABO06467 standard; protein; 907 AA.
XX
AC ABO06467;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human G-protein coupled receptor HG38.
XX
KW G-protein coupled receptor; GPCR; caudate nucleus related disorder;
KW neoplastic disorder; leukaemia; breast cancer; immune disorder; AIDS;
KW rheumatoid arthritis; neuronal disorder; Alzheimer's disease; receptor;
KW Parkinson's disease; respiratory disorder; bronchopulmonary disease;
KW pancreatic disease; ovarian disorder; ovarian carcinoma; colon disease;
KW Meig's syndrome; breast disease; thalamus related disorder; gene therapy;
KW amygdala related disorder; corpus callosum related disorder;
KW hippocampus related disorder.
XX
OS Homo sapiens.
XX
FN US2003027323-A1.
XX

PD 06-FEB-2003.

XX 26-SEP-2001; 2001US-00965536.

XX 27-SEP-2000; 2000US-0235832P.

PR 16-JAN-2001; 2001US-0261781P.

PR 19-JUL-2001; 2001US-0306605P.

PR 03-AUG-2001; 2001US-0310436P.

XX (FEDE/) FEDER J N.

XX (MINT/) MINTIER G.

XX (RAMA/) RAMANATHAN C S.

XX (HAWK/) HAWKEN D R.

XX Feder JN, Minter G, Ramanathan CS, Hawken DR;

XX WPI; 2003-466147/44.

XX New G-protein coupled receptor polypeptides, designated as HGPRMY5,
PT useful for preventing, treating or ameliorating a medical condition
PT related to the colon, breast, ovaries or immune system.

XX Disclosure; Fig 8; 93pp; English.

XX The invention relates to an isolated HGPRMY5 polypeptide. The
CC polypeptides, polynucleotides and methods are useful for preventing,
CC treating or ameliorating a medical condition such as a neoplastic
CC disorder e.g. leukaemia and breast cancer; immune disorder e.g. AIDS and
CC rheumatoid arthritis; neuronal disorder e.g. Alzheimer's disease and
CC Parkinson's disease; respiratory disorder e.g. bronchopulmonary disease
CC and pancreatic disease; ovarian disorder e.g. ovarian carcinoma and
CC Meig's syndrome; colon disease; breast disease; thalamus related disorder;
CC amygdala related disorder; corpus callosum related disorder; caudate
CC nucleus related disorder; hippocampus related disorder by administering
CC the GPCR polypeptide or its homologue. The present sequence represents
CC the amino acid sequence of a G-protein coupled receptor used to show
CC homology with the human G-protein coupled receptor, HGPRBMY

XX Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 6; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSRIGVLLSLPVLLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLRVDCSLGSEL 60

DB 1 MDTSRIGVLLSLPVLLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLRVDCSLGSEL 60

QY 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSKVLV 120

DB 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSKVLV 120

QY 121 LQNNQLRHVPTEALQNLRSQSLRLDANHSIYVPPSPCSGLHSLRHLWLDNALTEIPVQ 180

DB 121 LQNNQLRHVPTEALQNLRSQSLRLDANHSIYVPPSPCSGLHSLRHLWLDNALTEIPVQ 180

QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRHHISLGKKCFGLHSLETLD 240

DB 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRHHISLGKKCFGLHSLETLD 240

QY 241 LNYNNLDEFPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIDFVGRSA 300

DB 241 LNYNNLDEFPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIDFVGRSA 300

QY 301 FOHLPELRTLTNGASQITTEPDLTGTANLESRLTGTGAQISSLPQTVCNQLPNIQVLDLS 360

DB 301 FOHLPELRTLTNGASQITTEPDLTGTANLESRLTGTGAQISSLPQTVCNQLPNIQVLDLS 360

QY 361 YNLLEDLPSPVSVCKLOKIDLRHNEIYEIKVDTFQQLLSRLSLANWKIAIHPNAPST 420

DB 361 YNLLEDLPSPVSVCKLOKIDLRHNEIYEIKVDTFQQLLSRLSLANWKIAIHPNAPST 420

QY 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISENPFPELKIEMPVAYOCC 480


```
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISENPELKVEMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGDNSMDDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGDNSMDDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCHEHLLDGLMIRIGVWTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCHEHLLDGLMIRIGVWTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSPARHGAWNGVGVCHVIGPLSIPASESSVFLTLTAAAL 660
Db 601 AVNMLTGVSVAVLGADFTGSPARHGAWNGVGVCHVIGPLSIPASESSVFLTLTAAAL 660
QY 661 ERGFSVKYSAPETKAPFSSKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAPETKAPFSSKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGEP 720
QY 721 STMGWVVALIILNSLCFLMTTATYKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGWVVALIILNSLCFLMTTATYKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFTISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTVV 840
Db 781 PVAFLSFSSLINLTFTISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTVV 840
QY 841 WTRSKHPSLMSINSDVEKQSCDSTQALVFTFTSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVEKQSCDSTQALVFTFTSSITVDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPECL 907
Db 901 VAFVPECL 907

RESULT 5
ID ADC22783
AC ADC22783 standard; protein; 907 AA.
XX ADC22783;
XX 18-DEC-2003 (first entry)
DT Human G protein-coupled receptor (GPCR) polypeptide #39.
DE Human; G protein-coupled receptor; GPCR; transmembrane-6 region; TM6;
KW intracellular-3 region; IC3; receptor.
XX Homo sapiens.
XX US6555339-B1.
XX 29-APR-2003.
XX 13-OCT-1998; 98US-00170496.
XX 14-APR-1997; 97US-00839449.
XX 14-APR-1998; 98US-00060188.
XX 26-JUN-1998; 98US-0090783P.
XX 07-AUG-1998; 98US-0095677P.
XX (AREN-) ARENA PHARM INC.
XX Liaw CW, Behan DP, Chalmers DT;
XX WPI; 2003-742861/70.
XX N-PSDB; ADC22782.
XX Creating a constitutively active version of an endogenous human G protein
PT coupled receptor (GPCR) comprises substituting a specific amino acid in
PT the transmembrane-6 region with a different amino acid, and testing for
PT constitutive activity.
```

```
XX Example 1; SEQ ID NO 264; 221pp; English.
XX The invention relates to a method for treating a non-endogenous,
CC constitutively active version of an endogenous human G protein-coupled
CC receptor (GPCR) that has a transmembrane-6 (TM6) region and an
CC intracellular-3 (IC3) region, by substituting a specific amino acid in
CC the TM6 region with a different amino acid, and testing for constitutive
CC activity. The method is useful for creating a constitutively active
CC version of an endogenous human GPCR that comprises a transmembrane 6
CC region and an intracellular loop 3 region. The altered human GPCR
CC polypeptides are useful for screening test compounds for identification
CC of inverse agonists or partial agonists of GPCR polypeptides, which may
CC have therapeutic uses. The altered GPCRs may also be used in vivo or in
CC vitro in biological research. A nucleic acid encoding the altered GPCR
CC may be used to create a transgenic animal expressing the altered GPCR.
CC The method allows screening for compounds that modulate the activity of a
CC human G protein-coupled receptor without the need for provision of a
CC ligand for the receptor. This is particularly useful in allowing
CC screening of compounds against orphan receptors for which no ligand is
CC currently known. This sequence represents a human GPCR polypeptide of the
CC invention.
XX SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSRGLVLLSLPVLLQLATGSSPSRSGVLGRGCPHCHCEPDGRMLLRVDCSDLGLSEL 60
Db 1 MDTSRGLVLLSLPVLLQLATGSSPSRSGVLGRGCPHCHCEPDGRMLLRVDCSDLGLSEL 60
QY 61 PSNLSVFTSYLDLSNMNISQLLPNPLPSRLFEELRLAGNALTYIPKGAFTGLYSKVLIM 120
Db 61 PSNLSVFTSYLDLSNMNISQLLPNPLPSRLFEELRLAGNALTYIPKGAFTGLYSKVLIM 120
QY 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHI SYVPPSCFSGLSHLRHLMDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHI SYVPPSCFSGLSHLRHLMDNALTEIPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLHLHNNRHSIGKCFDGLHSLTLD 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLHLHNNRHSIGKCFDGLHSLTLD 240
QY 241 LNYNNLDEFPPTAIRTLNLKELGFHSNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
Db 241 LNYNNLDEFPPTAIRTLNLKELGFHSNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
QY 301 FOHLPELRTLINGASQITEFPDLTGTTANLSLTLTGAQISLSPOTVCNQLPNQLVDLS 360
Db 301 FOHLPELRTLINGASQITEFPDLTGTTANLSLTLTGAQISLSPOTVCNQLPNQLVDLS 360
QY 361 YNLLEDLPSFVSCQKLQKIDLRHNEIYEIKVDYTFQQLSLSLANLAWNKIATIHNAFST 420
Db 361 YNLLEDLPSFVSCQKLQKIDLRHNEIYEIKVDYTFQQLSLSLANLAWNKIATIHNAFST 420
QY 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISENPELKVEMPAYQCC 480
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISENPELKVEMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGDNSMDDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGDNSMDDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCHEHLLDGLMIRIGVWTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCHEHLLDGLMIRIGVWTIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSPARHGAWNGVGVCHVIGPLSIPASESSVFLTLTAAAL 660
Db 601 AVNMLTGVSVAVLGADFTGSPARHGAWNGVGVCHVIGPLSIPASESSVFLTLTAAAL 660
```


Qy 661 ERGFSVKYSAKFETKAPSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db |||||
Qy 661 ERGFSVKYSAKFETKAPSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db |||||
Qy 721 STMGVMVALIILNSLCFLMWTIAATKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
Db |||||
Qy 721 STMGVMVALIILNSLCFLMWTIAATKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
Db |||||
Qy 781 PVAFLSSSLINLTFISPEVTKFILLVVVPLPACILNPLLYILFPHFKEDILVSRKQTYV 840
Db |||||
Qy 781 PVAFLSSSLINLTFISPEVTKFILLVVVPLPACILNPLLYILFPHFKEDILVSRKQTYV 840
Db |||||
Qy 841 WTRSKHPSLMSINDVDEKQSCDSTQALVVTFTSSITVDLPSSVPSPAYPVTTESCHLSS 900
Db |||||
Qy 841 WTRSKHPSLMSINDVDEKQSCDSTQALVVTFTSSITVDLPSSVPSPAYPVTTESCHLSS 900
Db |||||
Qy 901 VAFVPCPL 907
Db |||||
Qy 901 VAFVPCPL 907
Db |||||
RESULT 6
ADE59150
ID ADE59150 standard; protein; 907 AA.
XX
AC ADE59150;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein O75473, SEQ ID NO 5041.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
WPI; 2003-268312/26.
DR GENBANK; O75473.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDTSLRGVLLSLPVLQLATGSGSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db |||||
Qy 1 MDTSLRGVLLSLPVLQLATGSGSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db |||||
Qy 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLV 120
Db |||||
Qy 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLV 120
Db |||||
Qy 121 LQNNQLRHVPTALQNLRLQSLRLDANHSYVPPSCFSGLSLRLHLDLWDDNALTEIPVQ 180
Db |||||
Qy 121 LQNNQLRHVPTALQNLRLQSLRLDANHSYVPPSCFSGLSLRLHLDLWDDNALTEIPVQ 180
Db |||||
Qy 181 AFRSLALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLETLD 240
Db |||||
Qy 181 AFRSLALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLETLD 240
Db |||||
Qy 241 LNYNNLDEFPTAIRTLSNLKELGPHSNIRSIPEKAFVGNPSLTIHFYNDPIQFVGSA 300
Db |||||
Qy 241 LNYNNLDEFPTAIRTLSNLKELGPHSNIRSIPEKAFVGNPSLTIHFYNDPIQFVGSA 300
Db |||||
Qy 301 FOHLPELRTLTNGASQITFPPDLTGANLESFLTGAQISSLPQTVCNQLPVLQVLDLS 360
Db |||||
Qy 301 FOHLPELRTLTNGASQITFPPDLTGANLESFLTGAQISSLPQTVCNQLPVLQVLDLS 360
Db |||||
Qy 361 YNLLEDLPSPFVSVCKQKIDLRHNEIYEIKVDYTFQQLLSRLSLANWKTAIHPNAPST 420
Db |||||
Qy 361 YNLLEDLPSPFVSVCKQKIDLRHNEIYEIKVDYTFQQLLSRLSLANWKTAIHPNAPST 420
Db |||||
Qy 421 LPSLIKLDLSSNLLSPFITGLHGLTHLKLGTGNHALQSLISSENFPELKVEMPVAYOCC 480
Db |||||
Qy 421 LPSLIKLDLSSNLLSPFITGLHGLTHLKLGTGNHALQSLISSENFPELKVEMPVAYOCC 480
Db |||||
Qy 481 AFGVCENAYKISNOWNKDSSMDLHKDAGMFQADDERDLEDLDFEEDLKALHSVQ 540
Db |||||
Qy 481 AFGVCENAYKISNOWNKDSSMDLHKDAGMFQADDERDLEDLDFEEDLKALHSVQ 540
Db |||||
Qy 541 CSPSPGPFKPCHEHLLDGWLIRIGVWTIAVALTALCNALVTSTVFRSPLYIYSPKLLIGVIA 600
Db |||||
Qy 541 CSPSPGPFKPCHEHLLDGWLIRIGVWTIAVALTALCNALVTSTVFRSPLYIYSPKLLIGVIA 600
Db |||||
Qy 601 AVNNLTGVSSAVLAGVDAFTFGSFARHAWWENGVGCHVIGFLSIFASESVFLLTAAAL 660
Db |||||
Qy 601 AVNNLTGVSSAVLAGVDAFTFGSFARHAWWENGVGCHVIGFLSIFASESVFLLTAAAL 660
Db |||||
Qy 661 ERGFSVKYSAKFETKAPSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db |||||
Qy 661 ERGFSVKYSAKFETKAPSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db |||||
Qy 721 STMGVMVALIILNSLCFLMWTIAATKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
Db |||||
Qy 721 STMGVMVALIILNSLCFLMWTIAATKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
Db |||||

QY 781 PVAFSLFSSLINLTPISEVIFKIFILLVVVPLPACLNPLLYILFNPHEKEDLVSLKQTYV 840
 Db 781 PVAFSLFSSLINLTPISEVIFKIFILLVVVPLPACLNPLLYILFNPHEKEDLVSLKQTYV 840
 QY 841 WTRSKHPSLMSINSDVKEQSCDSTOALVTFSSITVDLPSSVPSPAYPVTSCHLSS 900
 Db 841 WTRSKHPSLMSINSDVKEQSCDSTOALVTFSSITVDLPSSVPSPAYPVTSCHLSS 900
 QY 901 VAFVPC 907
 Db 901 VAFVPC 907
 RESULT 7
 ADE59153
 ID ADE59153 standard; protein; 907 AA.
 XX
 AC ADE59153;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein O75473, SEQ ID NO 5044.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 FN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEMO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 DR WPI; 2003-268312/26.
 DR GENBANK; O75473.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PS preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page: 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the patented
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 907 AA;
 Query Match 100.0%; Score 4702; DB 7; Length 907;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRLMRVDCSDGLSEL 60
 Db 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRLMRVDCSDGLSEL 60
 QY 61 PSNLSVFTSYLDLSNNNISQLLPNPLSLRLEELRAGNALTYIPKGAFTGLYSKVLV 120
 Db 61 PSNLSVFTSYLDLSNNNISQLLPNPLSLRLEELRAGNALTYIPKGAFTGLYSKVLV 120
 QY 121 LQNNQLRHVPTEALQNLRLSLQSLRLDANHI SYVPPSCFSGLSLRLHLDNNAITFIPVQ 180
 Db 121 LQNNQLRHVPTEALQNLRLSLQSLRLDANHI SYVPPSCFSGLSLRLHLDNNAITFIPVQ 180
 QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLLHNNRHSLSGKFCFGLHSLFTLD 240
 Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLLHNNRHSLSGKFCFGLHSLFTLD 240
 QY 241 LNYNNLDEFPTAIRTLNKLGFHNSNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
 Db 241 LNYNNLDEFPTAIRTLNKLGFHNSNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
 QY 301 FOHLPELRTLINGASQITTEPDLTGTAANLSSLTGTGAQISSLPOTVCNQLPQLVDLS 360
 Db 301 FOHLPELRTLINGASQITTEPDLTGTAANLSSLTGTGAQISSLPOTVCNQLPQLVDLS 360
 QY 361 YNLLEDLPFSVSCQKQKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIAIHHNAPST 420
 Db 361 YNLLEDLPFSVSCQKQKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIAIHHNAPST 420
 QY 421 LPSLIKLDLSNLSLSPFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPYAYOC 480
 Db 421 LPSLIKLDLSNLSLSPFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPYAYOC 480
 QY 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAODERLEDFLLDFEEDLKALHSVQ 540
 Db 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAODERLEDFLLDFEEDLKALHSVQ 540
 QY 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
 Db 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
 QY 601 AVNMLTGVSNAVLGADAFTEGSPARHAGWENGCHVIGFLSTIFASESSVFLITLAL 660
 Db 601 AVNMLTGVSNAVLGADAFTEGSPARHAGWENGCHVIGFLSTIFASESSVFLITLAL 660
 QY 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALLATWAAVPLLGSKYKASPLCLPFGEP 720
 Db 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALLATWAAVPLLGSKYKASPLCLPFGEP 720
 QY 721 STMGYMWALILLNSLCLFLMWTIATYKLYCNLDKGLNIDWCSMKVHIALLLFTNCILNC 780
 Db 721 STMGYMWALILLNSLCLFLMWTIATYKLYCNLDKGLNIDWCSMKVHIALLLFTNCILNC 780
 QY 781 PVAFSLFSSLINLTPISEVIFKIFILLVVVPLPACLNPLLYILFNPHEKEDLVSLKQTYV 840
 Db 781 PVAFSLFSSLINLTPISEVIFKIFILLVVVPLPACLNPLLYILFNPHEKEDLVSLKQTYV 840
 QY 841 WTRSKHPSLMSINSDVKEQSCDSTOALVTFSSITVDLPSSVPSPAYPVTSCHLSS 900
 Db 841 WTRSKHPSLMSINSDVKEQSCDSTOALVTFSSITVDLPSSVPSPAYPVTSCHLSS 900


```
PR 13-OCT-1998; 98US-00170496.
XX (LIAW/) LIAW C W.
PA (BEHA/) BEHAN D P.
PA (CHAL/) CHALMERS D T.
XX
PI Liaw CW, Behan DP, Chalmers DT;
XX
DR WPI; 2003-801247/75.
DR N-PSDB; ADH14255.
XX
PT New constitutively active, non-endogenous version of an endogenous human
PT G protein-coupled receptor for the identification of therapeutic
PT compounds, such as agonists.
XX
PS Example 1; SEQ ID NO 264; 227bp; English.
XX
CC The invention relates to a constitutively active, non-endogenous version
CC of an endogenous human G protein-coupled receptor (GPCR). The GPCR is
CC used for screening therapeutic compounds as inverse agonists, agonists or
CC partial agonists. The GPCR can be also be used to elucidate and
CC understand the roles of GPCRs in normal and diseased humans. The GPCR
CC need not be purified and isolated to be used to screen for therapeutic
CC compounds. The utility of the GPCR as a research tool is enhanced because
CC the role of a particular receptor can be understood before the endogenous
CC ligand is identified. The present sequence is used in the exemplification
CC of the present invention.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLGVLISLPVLLQATGSSPRSGVLRGCPHCHCEPDGRLARVDCSDLGSEL 60
DB 1 MDTSLGVLISLPVLLQATGSSPRSGVLRGCPHCHCEPDGRLARVDCSDLGSEL 60
QY 61 PSLNSVFTSYLDLNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLV 120
DB 61 PSLNSVFTSYLDLNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSQSLRDANHISVVPSCFSGLSRLHRLWLDNALTEIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRSQSLRDANHISVVPSCFSGLSRLHRLWLDNALTEIPVQ 180
QY 181 APRSLAQAMTALNKHIIHIDYAFGNLSLVLHLHNNRIHSLGKCFDGLHSLDTLD 240
DB 181 APRSLAQAMTALNKHIIHIDYAFGNLSLVLHLHNNRIHSLGKCFDGLHSLDTLD 240
QY 241 LNNYLNDEFPPTAIRLSNLKELGFHNNIRSIPEKAFVGNPSLIITHFYDNPQFVGRSA 300
DB 241 LNNYLNDEFPPTAIRLSNLKELGFHNNIRSIPEKAFVGNPSLIITHFYDNPQFVGRSA 300
QY 301 FOHLPELRTLTLNGASQITFEPLDTGTANLESLLTGAQISSLPQVCNQLPNQLVLDLS 360
DB 301 FOHLPELRTLTLNGASQITFEPLDTGTANLESLLTGAQISSLPQVCNQLPNQLVLDLS 360
QY 361 YNLLEDLPSFSVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIALIHNAFST 420
DB 361 YNLLEDLPSFSVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIALIHNAFST 420
QY 421 LPSLTKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIMPYAYOCC 480
DB 421 LPSLTKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIMPYAYOCC 480
QY 481 AFGVCENAYKISQNWKNQDNSSMDLHKDAGMFOAQDERDLEDLFDLFEEDLKALHSVQ 540
DB 481 AFGVCENAYKISQNWKNQDNSSMDLHKDAGMFOAQDERDLEDLFDLFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCBHLDDGMLIRIGWVTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
DB 541 CSPSPGPKPCBHLDDGMLIRIGWVTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
```

```
QY 601 AVNMLTGVSASVLAGVDATFTGSPARHGAWWENGVCVIGFLSIPASESSVFLTLAAL 660
DB 601 AVNMLTGVSASVLAGVDATFTGSPARHGAWWENGVCVIGFLSIPASESSVFLTLAAL 660
QY 661 ERGFSVKYSKAFETKAPESLSKVILLCALLALTMAAVPLLGSGSKYGASPLCLPFGEP 720
DB 661 ERGFSVKYSKAFETKAPESLSKVILLCALLALTMAAVPLLGSGSKYGASPLCLPFGEP 720
QY 721 STMGYVALIILNLSLCFLMMTIAYTKLYCNLDKGDLENIDWDCSMVKHIALLLFTNCILNC 780
DB 721 STMGYVALIILNLSLCFLMMTIAYTKLYCNLDKGDLENIDWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFSLFSSSLNLTFTISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTVV 840
DB 781 PVAFSLFSSSLNLTFTISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTVV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITYDLPSSVPSPAYPVTESCHLSS 900
DB 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITYDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907
RESULT 10
ADN40013
ID ADN40013 standard; protein; 907 AA.
XX
AC ADN40013;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C383.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerability; Gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-039775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX
```


PI Afaf D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DR, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
XX WPI; 2003-468649/44.
XX N-PSDB; ADN39796.
XX

PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.

PS Claim 12; SEQ ID NO C383; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.

XX Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLLQATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSLGSEL 60
Db 1 MDTSLRGVLLSLPVLLQATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSLGSEL 60

Qy 61 PSNLSVFTSYLDLSMNNISQLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKVLV 120
Db 61 PSNLSVFTSYLDLSMNNISQLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKVLV 120

Qy 121 LQNNQLRHVPTEALQNLRSQSLRLDANHI SYVPPSCFSLHSLRHLWLDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSQSLRLDANHI SYVPPSCFSLHSLRHLWLDNALTEIPVQ 180

Qy 181 AFRSLAQMTALANKTHHPDYAFGNLSLVLHLHNNRIHSLGKKCFDGLHSLETLD 240
Db 181 AFRSLAQMTALANKTHHPDYAFGNLSLVLHLHNNRIHSLGKKCFDGLHSLETLD 240

Qy 241 LNYNNLDEFPPTAIRTLNSLKLGHFSNNIRSIPEKAFVGNPSLTIHFYDNPQIFVGRSA 300
Db 241 LNYNNLDEFPPTAIRTLNSLKLGHFSNNIRSIPEKAFVGNPSLTIHFYDNPQIFVGRSA 300

Qy 301 FOHLPELRTLNGASQITTEPDLTGNTANLESILTGAQISSLPQTVCNQPNLQVLDLS 360
Db 301 FOHLPELRTLNGASQITTEPDLTGNTANLESILTGAQISSLPQTVCNQPNLQVLDLS 360

Qy 361 YNLLEDLPFSVCQKQKIDLRHNEIYKIDVTFOQLLSLRLSLNLANWKIAIHPNAFST 420
Db 361 YNLLEDLPFSVCQKQKIDLRHNEIYKIDVTFOQLLSLRLSLNLANWKIAIHPNAFST 420

Qy 421 LPSLTKLDLSNLSLSPFITGLHGLTHLTKGNHALQSLISSENPPELKIEMPYAYOC 480
Db 421 LPSLTKLDLSNLSLSPFITGLHGLTHLTKGNHALQSLISSENPPELKIEMPYAYOC 480

Qy 481 AFGVCENAYKISQNWKNQDNSSMDLHKDKAGMQAODERDLEFLDFFEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNQDNSSMDLHKDKAGMQAODERDLEFLDFFEDLKALHSVQ 540

Qy 541 CSPSPGPFKCEHLLDGWLRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Db 541 CSPSPGPFKCEHLLDGWLRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Qy 601 AVNMLTGSSAVLAGVDAFTFGSPARHGAWWENGCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGSSAVLAGVDAFTFGSPARHGAWWENGCHVIGFLSIFASESSVFLTLAAL 660

Qy 661 ERGFSVKYSKAFETKAPFSSIKVILLCALLATMAAVPLLGSKYKASPLCLPFGEP 720
Db 661 ERGFSVKYSKAFETKAPFSSIKVILLCALLATMAAVPLLGSKYKASPLCLPFGEP 720

Qy 721 STMGYVALIILNLSLCLFLMWTIATKLYCNLDKGDLENIDCSMVGHIALLLFNCLINC 780
Db 721 STMGYVALIILNLSLCLFLMWTIATKLYCNLDKGDLENIDCSMVGHIALLLFNCLINC 780

Qy 781 PVAFLSFSSILNLTIFISPEVIKFIILLVVVPLPACINPLLYLILFNPHKEDLVSRKQTYV 840
Db 781 PVAFLSFSSILNLTIFISPEVIKFIILLVVVPLPACINPLLYLILFNPHKEDLVSRKQTYV 840

Qy 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVTESCHLSS 900

Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907

RESULT 11
ADN39531
ID ADN39531 standard; protein; 907 AA.
XX
AC ADN39531;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A131.
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PP 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX

PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI: 2003-468649/44.
DR N-PSDB; ADN39530.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO A131; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDTSLRGVLLSLPVLQATGSSPSRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSLRGVLLSLPVLQATGSSPSRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Qy 61 PSNLSVFTSYLDLNNISQLLPNPLPSLRFLELRAGNALTYPKGAPGLYSKVLV 120
Db 61 PSNLSVFTSYLDLNNISQLLPNPLPSLRFLELRAGNALTYPKGAPGLYSKVLV 120
Qy 121 LONNQLRHVPTEALQNLRSLOSLRDANHI SVPPSCFSGHLSRHLWLDNALTETPVQ 180
Db 121 LONNQLRHVPTEALQNLRSLOSLRDANHI SVPPSCFSGHLSRHLWLDNALTETPVQ 180
Qy 181 AFRSLALQAMTALNKHIIHPDYAFGNLSLVVLHNNRIHSLGKCFDGLHSLETLD 240
Db 181 AFRSLALQAMTALNKHIIHPDYAFGNLSLVVLHNNRIHSLGKCFDGLHSLETLD 240
Qy 241 LNNYNDLDEFTAIRTLNLSKELGHSNNIRSIPEKAFVGNPSLITIHFDYNDPIQVGRSA 300
Db 241 LNNYNDLDEFTAIRTLNLSKELGHSNNIRSIPEKAFVGNPSLITIHFDYNDPIQVGRSA 300
Qy 301 FOHLPELRTLNGASQITTEPDLTGNTANLESITLGTGAQISLPQVTCNQLPNLVLDLS 360
Db 301 FOHLPELRTLNGASQITTEPDLTGNTANLESITLGTGAQISLPQVTCNQLPNLVLDLS 360
Qy 361 YNLEDLPSFVSCQKQKIDLRHNEIYKVDYTFQOLLSLRLNLANWKIAIHPNAPST 420
Db 361 YNLEDLPSFVSCQKQKIDLRHNEIYKVDYTFQOLLSLRLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKDLSSNLSFFPITGLHGLTHLTKGNHALQSLISSENPPELKIEMPYAYQCC 480
Db 421 LPSLIKDLSSNLSFFPITGLHGLTHLTKGNHALQSLISSENPPELKIEMPYAYQCC 480
Qy 481 AFGVCENAYKISNQWKNKGNSSMDLHKDKAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Db 481 AFGVCENAYKISNQWKNKGNSSMDLHKDKAGMFOAQDERDLEFLDFFEDLKALHSVQ 540

Qy 541 CSPSPGPFKPCHEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLVISPIKLLIGVIA 600
Db 541 CSPSPGPFKPCHEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLVISPIKLLIGVIA 600
Qy 601 AVNMLTGVSAAVLAVDAFTGSGFARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSAAVLAVDAFTGSGFARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
Qy 661 ERGSVKYSKAFETKAPFSSILKVIILLCALILATMAAVPLLGSGSKYGASPLCLPFGEP 720
Db 661 ERGSVKYSKAFETKAPFSSILKVIILLCALILATMAAVPLLGSGSKYGASPLCLPFGEP 720
Qy 721 STMGYVALIILNSLCFLMWTIATKYCNLDKGLNINWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNSLCFLMWTIATKYCNLDKGLNINWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSFSSILNLPISPEVTKFILLVVPPLACINPLLYILFNPHKEDLVSLRKQTVV 840
Db 781 PVAFLSFSSILNLPISPEVTKFILLVVPPLACINPLLYILFNPHKEDLVSLRKQTVV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVTESCHLSS 900
Qy 901 VAFVPCCL 907
Db 901 VAFVPCCL 907
RESULT 12
ADN39628
ID ADN39628 standard; protein; 907 AA.
XX AC ADN39628;
XX DT 17-JUN-2004 (first entry)
XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A228.
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulneryary; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN WO2003042661-A2.
XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.

PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI; 2003-468649/44.
DR N-PSDB; ADN39627.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
XX Claim 12; SEQ ID NO A228; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
XX
SQ Sequence 907 AA;

Query Match 100.0%; Score 4702; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLQATCGSPRGVLLRGCPHCHCEPDGRMLLRVDCSLGSEL 60
Db 1 MDTSLRGVLLSLPVLQATCGSPRGVLLRGCPHCHCEPDGRMLLRVDCSLGSEL 60

Qy 61 PSNLVFTSYLDLSNNISQILNPLPSLRLFLRLAGNALTYPKGAFTGLYSKVLN 120
Db 61 PSNLVFTSYLDLSNNISQILNPLPSLRLFLRLAGNALTYPKGAFTGLYSKVLN 120

Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRDANHSYVPPSPFGLSHSLRLWLDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRDANHSYVPPSPFGLSHSLRLWLDNALTEIPVQ 180

Qy 181 AFRSLSALQAMTLAKNTHHPDYAFGNLSLVLLHNNRIHSLGRKCPDGLHSLTLD 240
Db 181 AFRSLSALQAMTLAKNTHHPDYAFGNLSLVLLHNNRIHSLGRKCPDGLHSLTLD 240

Qy 241 LNYNNLDEFPFAIRTLNKLGLFHSNNIRSIPEKAFVGNPSLTIHFYDNPFOVGRSA 300
Db 241 LNYNNLDEFPFAIRTLNKLGLFHSNNIRSIPEKAFVGNPSLTIHFYDNPFOVGRSA 300

Qy 301 FOHLPELRTLNGASQITEFPDITGTANLESRLTGAQISLSPQTVCNQLPNQLVDLS 360
Db 301 FOHLPELRTLNGASQITEFPDITGTANLESRLTGAQISLSPQTVCNQLPNQLVDLS 360

Qy 361 YNLLDLPFSVCQKQKIDLRHNEIYEIKVDYTFQQLLSLSLNLAWNKIAIHPNAPST 420
Db 361 YNLLDLPFSVCQKQKIDLRHNEIYEIKVDYTFQQLLSLSLNLAWNKIAIHPNAPST 420

Qy 421 LPSLIKLDLSNLLSSPFTIGLHGLTHLKTGNHALQSLISSENFPELKVIEPYAYQCC 480
Db 421 LPSLIKLDLSNLLSSPFTIGLHGLTHLKTGNHALQSLISSENFPELKVIEPYAYQCC 480

Qy 481 AFGVCENAYKISNQWNGDSSMDLHKKDAGMFQAQDERDLEDFLDFFEDLKALHSVQ 540

Db 481 AFGVCENAYKISNQWNGDSSMDLHKKDAGMFQAQDERDLEDFLDFFEDLKALHSVQ 540
Qy 541 CSPSPGPFKPCCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPFKPCCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGVSAAVLGADAFPGSFARHAGWENGVCCHVIGFLSIFASESVFLLTAAAL 660
Db 601 AVNMLTGVSAAVLGADAFPGSFARHAGWENGVCCHVIGFLSIFASESVFLLTAAAL 660
Qy 661 ERGFSVKYSKAFETKAPFSSKVLIIICALLATMAAVPLLGSGSKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSKAFETKAPFSSKVLIIICALLATMAAVPLLGSGSKYGASPLCLPFGEP 720
Qy 721 STMGYVALIILNLSLCLFLMMTIATKLYCNLDKGDLENINDCSMVKHIALLLFNCLINC 780
Db 721 STMGYVALIILNLSLCLFLMMTIATKLYCNLDKGDLENINDCSMVKHIALLLFNCLINC 780
Qy 781 PVALFSSSLINLTFFISPEVTKFLLVVVPLPACLNPLLYLLENPHFKEDLVSRKQTYV 840
Db 781 PVALFSSSLINLTFFISPEVTKFLLVVVPLPACLNPLLYLLENPHFKEDLVSRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVFTSSSITYDLPSSVPSPAYPVTESCHLS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVFTSSSITYDLPSSVPSPAYPVTESCHLS 900
Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907

RESULT 13
ADO29408
ID ADO29408 standard; protein; 907 AA.
XX AC ADO29408;
XX AC ADO29408;
XX DT 29-JUL-2004 (first entry)
XX DE Human GPCR GPR49, SEQ ID NO:510.
XX DE
XX KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antimanic;
KW cytosolic; antinflammatory; vasotropic; antianginal; antiarrhythmic;
KW CNS; central nervous system; respiratory; antiarrhoeic; antidiabetic;
KW virucide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW dermatological; antiulcer; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW receptor.
XX OS Homo sapiens.
XX OS
XX WO2004040000-A2.
XX PN
XX PD 13-MAY-2004.
XX PD
XX PF 09-SEP-2003; 2003WO-US028226.
XX PF
XX PR 09-SEP-2002; 2002US-0409303P.
XX PR
XX PR 09-APR-2003; 2003US-0461329P.
XX PR
XX (PRIM-) PRIMAL INC.
XX FA
XX PI Gaitanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F;
XX PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;

XX WPI; 2004-390329/36.
DR N-PSDB; AD029923.
XX
PT Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
PS Claim 151; SEQ ID NO 510; 542bp; English.
XX
CC The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridize to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
CC invention. Note: The full sequence data for this patent did not form part
CC of the printed specification; those sequences not shown were obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 907 AA;

Query Match
Best Local Similarity 100.0%; Score 4702; DB 8; Length 907;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

QY 61 PNLSVFTSYLDLNMNLSQLLPNLPSLRFLEELRAGNALTYTPKGAFTGLYSKVLV 120
DB 61 PNLSVFTSYLDLNMNLSQLLPNLPSLRFLEELRAGNALTYTPKGAFTGLYSKVLV 120

QY 121 LQNNQLRHVPTEALQNLRLSLQSLRDANHSVVPSPCSGLSLRHLMDNALTEIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRLSLQSLRDANHSVVPSPCSGLSLRHLMDNALTEIPVQ 180

QY 181 AFRSLALQAMTLALNKIHHIDYAFGNLSLVLHLHNNRHSGLKKFCGLHSLETID 240
DB 181 AFRSLALQAMTLALNKIHHIDYAFGNLSLVLHLHNNRHSGLKKFCGLHSLETID 240

QY 241 LNNYNNLDEPPTAIRLSNKLGFHNSNIRSIPEKAFVGNPSLTIHFYDNPQVGRSA 300
DB 241 LNNYNNLDEPPTAIRLSNKLGFHNSNIRSIPEKAFVGNPSLTIHFYDNPQVGRSA 300

QY 301 FOHLPELRLTLNGASQITEFPDLTGTLNLSLITGAQISSLPQVNCNQLNQLVDLS 360
DB 301 FOHLPELRLTLNGASQITEFPDLTGTLNLSLITGAQISSLPQVNCNQLNQLVDLS 360

QY 361 YNLEDLPSFSVCOKIQKIDLRHNEIYEIKVDYTFQQLSLSLANWKIALIHNAPST 420
DB 361 YNLEDLPSFSVCOKIQKIDLRHNEIYEIKVDYTFQQLSLSLANWKIALIHNAPST 420

QY 421 LPSLIKLDLSSNLSLSSPFTITGLHGLTHLKLATGNHALQSLISSENPPELKIEMPIYQCC 480
DB 421 LPSLIKLDLSSNLSLSSPFTITGLHGLTHLKLATGNHALQSLISSENPPELKIEMPIYQCC 480

QY 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDLDFEEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEDLDFEEDLKALHSVQ 540

QY 541 CSPSPGPKPECHLLDGLWLIIRIGVWTIAVLTATCNALVTSTVFSPLYSPIKLLIGVIA 600
DB 541 CSPSPGPKPECHLLDGLWLIIRIGVWTIAVLTATCNALVTSTVFSPLYSPIKLLIGVIA 600

QY 601 AVNMLTGVSASVAVAGDAFTGSPARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSASVAVAGDAFTGSPARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660

QY 661 ERGFSVKYSAKFETKAPSSKLVIIICALLALTAALTAAPVLLGGSKYGASPLCLPFGEP 720
DB 661 ERGFSVKYSAKFETKAPSSKLVIIICALLALTAALTAAPVLLGGSKYGASPLCLPFGEP 720

QY 721 STMGYVVALIILNSLCFLMMTIATKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
DB 721 STMGYVVALIILNSLCFLMMTIATKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780

QY 781 PVAFLSPSSLINLTFISPEVIFKILLVVVPLPACLNPLLYILFNPHFKEDVSLRKQTVV 840
DB 781 PVAFLSPSSLINLTFISPEVIFKILLVVVPLPACLNPLLYILFNPHFKEDVSLRKQTVV 840

QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVVTESCHLSS 900
DB 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVVTESCHLSS 900

QY 901 VAFVPC 907
DB 901 VAFVPC 907

RESULT 14
ADQ80369
ID ADQ80369 standard; protein; 907 AA.
AC ADQ80369;
XX
XX 21-OCT-2004 (first entry)
DE G protein-coupled receptor 49 protein;
KW Cytostatic; epidermal-growth-factor-receptor modulator; identification;
KW therapeutic response; cancer; EGFR; biomarker.
XX Homo sapiens.
XX WO2004063709-A2.
XX 29-JUL-2004.
XX 08-JAN-2004; 2004WO-US0000368.
XX 08-JAN-2003; 2003US-0438735P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Amler LC, Januario-T;
XX WPI; 2004-544114/52.
XX N-PSDB; ADQ80249.
XX
XX Identifying a mammal that will respond therapeutically to a method of
PT treating cancer comprises the level of a biomarker in a mammal

PT before and after exposure to an epidermal growth factor receptor (EGFR)
XX modulator.
PS Disclosure; SEQ ID NO 141; 520pp; English.
XX
CC The invention relates to a method of identifying a mammal that will
CC respond therapeutically to a method of treating cancer by administering
CC an epidermal growth factor receptor (EGFR) modulator by comparing the
CC level of a biomarker in a mammal before and after exposure to an EGFR
CC modulator. The method comprises: (a) measuring, in the mammal, the level
CC of at least one biomarker identified in the specification; (b) exposing
CC the mammal to the EGFR modulator; and (c) measuring in the mammal the
CC level of the biomarker, where a difference in the level in step (c)
CC compared to step (a) indicates that the mammal will respond
CC therapeutically to the method of treating cancer. The method and
CC biomarkers are useful for identifying a mammal that will respond
CC therapeutically to a method of treating cancer by administering an
CC epidermal growth factor receptor (EGFR) modulator. This sequence
CC corresponds to one of the biomarkers whose levels of expression is
CC measured in the method of the invention.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 4702; DB 8; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLRLGVLISLPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGISEL 60
DB 1 MDTSLRLGVLISLPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGISEL 60
QY 61 PSNLSVFTSYDLSNNISQLLPPLSLRFLERLGNALTYIPKGAFTGLSKVLJM 120
DB 61 PSNLSVFTSYDLSNNISQLLPPLSLRFLERLGNALTYIPKGAFTGLSKVLJM 120
QY 121 LQNNQLRHVPTEALQNLRLSLQSLRLDANHSYVPPSCFSGHLSRLHLWLDNALTEIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRLSLQSLRLDANHSYVPPSCFSGHLSRLHLWLDNALTEIPVQ 180
QY 181 AFRSLALQMTALANKIHHIPDYAFGNLSLVLLHNNRIHSLGKCKPGLHSLTLD 240
DB 181 AFRSLALQMTALANKIHHIPDYAFGNLSLVLLHNNRIHSLGKCKPGLHSLTLD 240
QY 241 LNYNLDLDFPTAIRLSNLKELGHSNNIRSIPEKAFVGNPSLTIHFYDNPFOVGRSA 300
DB 241 LNYNLDLDFPTAIRLSNLKELGHSNNIRSIPEKAFVGNPSLTIHFYDNPFOVGRSA 300
QY 301 FOHLPELRTLINGASQITTEPDLTGTLNLSLTLTGAQISLPTQVCNQLPNLQVLDLS 360
DB 301 FOHLPELRTLINGASQITTEPDLTGTLNLSLTLTGAQISLPTQVCNQLPNLQVLDLS 360
QY 361 YNLLEDLPFSVCOKLQIDLRHNEIYIKVDVTFQQLLSLSRLMANNKIAIHPNAPST 420
DB 361 YNLLEDLPFSVCOKLQIDLRHNEIYIKVDVTFQQLLSLSRLMANNKIAIHPNAPST 420
QY 421 LPSLIKLDLSNLSLSSPITGLHGLTHLKGHALQSLISSENFPELKVEMPYAVOC 480
DB 421 LPSLIKLDLSNLSLSSPITGLHGLTHLKGHALQSLISSENFPELKVEMPYAVOC 480
QY 481 AFGVCENAYKISNOWNKDNSSMDLHKDKAGMFOQDERDLEFLDFFEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKDNSSMDLHKDKAGMFOQDERDLEFLDFFEDLKALHSVQ 540
QY 541 CSPSPGPKPEHLLDGLHRIIGVWTTAVLALTCNALVTSTVFRSPLYSPKLLIGVIA 600
DB 541 CSPSPGPKPEHLLDGLHRIIGVWTTAVLALTCNALVTSTVFRSPLYSPKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSGFARHGAWENGVCVIGFLSTFASSESVLLTAAAL 660
DB 601 AVNMLTGVSVAVLGADFTGSGFARHGAWENGVCVIGFLSTFASSESVLLTAAAL 660
QY 661 ERGFSVKYSAKFETKAPFSSLKVILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720

DB 661 ERGFSVKYSAKFETKAPFSSLKVILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720
QY 721 STMGYVALILLNSLCFLMMTTIATKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
DB 721 STMGYVALILLNSLCFLMMTTIATKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFISPEVIKFIILLVVPPLPACINPLLYLILFNPHFKEDLVSLRKQTYV 840
DB 781 PVAFLSFSSLINLTFISPEVIKFIILLVVPPLPACINPLLYLILFNPHFKEDLVSLRKQTYV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSTSSITVDLPSPSPVPAYPVTESCHLSS 900
DB 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSTSSITVDLPSPSPVPAYPVTESCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907
RESULT 15
ADR67868
ID ADR67868 standard; protein; 907 AA.
XX
XX ADR67868;
AC
XX 18-NOV-2004 (first entry)
DT
XX Human HG38 protein.
DE
XX human; G protein-coupled receptor; GPCR; HG38; colon; lung; cancer.
KW
XX Homo sapiens.
OS
XX WO2004074436-A2.
FN
XX 02-SEP-2004.
PD
XX 11-FEB-2004; 2004WO-US004060.
PF
XX 19-FEB-2003; 2003US-0448959P.
PR
XX (INCY-) INCYTE CORP.
PA
XX Lasek AW;
PI
XX WPI; 2004-652946/63.
DR
XX N-ESDB; ADR67869.
XX
XX Detecting colon or lung cancer, by determining amount of protein in
XX sample, comparing amount of protein to standard, and differential
XX expression of protein in sample indicates colon or lung cancer.
XX
XX Claim 4; SEQ ID NO 1; 79pp; English.
XX
XX This sequence represents the human G protein-coupled receptor (GPCR)
XX known as HG38. The HG38 protein and corresponding nucleic acid, may be
XX used in the method of the invention for detecting colon or lung cancer.
XX The first method involves performing an assay to determine the amount of
XX HG38 in a sample of colon or lung tissue, and comparing the amount of
XX protein to standard, thus detecting expression of protein in sample,
XX where differential expression of protein in sample when compared with the
XX standard is diagnostic of colon or lung cancer. The second method
XX involves hybridizing a composition comprising the HG38 coding sequence,
XX or its complement, and a labelling moiety, to nucleic acids of a sample
XX of colon or lung tissue under conditions to form at least one
XX hybridization complex, detecting hybridization complex formation, and
XX comparing complex formation of the polynucleotide in the sample relative to
XX differential expression of the polynucleotide in the sample relative to
XX the standard and is diagnostic of a colon or lung cancer. This first
XX method enables earlier diagnosis before the patient is symptomatic. Anti-
XX HG38 antibodies are useful for treating colon or lung cancer.
XX
XX Sequence 907 AA;

Query Match		100.0%;	Score 4702;	DB 8;	Length 907;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches	907;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1	MDTSRLGVLLSPVLLQATGSSPRSGVLLRGCPCHCEPDGRMLLRVDCSDGLSEL	60		
Db	1	MDTSRLGVLLSPVLLQATGSSPRSGVLLRGCPCHCEPDGRMLLRVDCSDGLSEL	60		
Qy	61	PSNLSVFTSYLDLNMNLSQILLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSKVLV	120		
Db	61	PSNLSVFTSYLDLNMNLSQILLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSKVLV	120		
Qy	121	LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSRLHMLDDNALTEIPVQ	180		
Db	121	LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSRLHMLDDNALTEIPVQ	180		
Qy	181	AFRSLAQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLTLD	240		
Db	181	AFRSLAQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLTLD	240		
Qy	241	LYNNLDEFPPTAIRLSNLKELGFHNSNIRSIPEKAFVGNPSLIITHFYDNPFIQVGRSA	300		
Db	241	LYNNLDEFPPTAIRLSNLKELGFHNSNIRSIPEKAFVGNPSLIITHFYDNPFIQVGRSA	300		
Qy	301	FOHLPRLTLTLNGASQITEFPDLTGTANLESILTGTAGISSLPQTVCNQLPNQLVLDLS	360		
Db	301	FOHLPRLTLTLNGASQITEFPDLTGTANLESILTGTAGISSLPQTVCNQLPNQLVLDLS	360		
Qy	361	YHLEDLPSFSVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIAIHHNAPST	420		
Db	361	YHLEDLPSFSVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIAIHHNAPST	420		
Qy	421	LPSLIKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEMPYAYOCC	480		
Db	421	LPSLIKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEMPYAYOCC	480		
Qy	481	AFGVCENAYKISNQNKNGNSMDDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ	540		
Db	481	AFGVCENAYKISNQNKNGNSMDDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ	540		
Qy	541	CSPSPGPKPCBHLDDGMLIRIGVWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA	600		
Db	541	CSPSPGPKPCBHLDDGMLIRIGVWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA	600		
Qy	601	AVNMLTGVSVAVLGADFTGSGFARHGAWENGCHVIGFELSI FASESSVFLTLAAL	660		
Db	601	AVNMLTGVSVAVLGADFTGSGFARHGAWENGCHVIGFELSI FASESSVFLTLAAL	660		
Qy	661	ERGSVKYSAPETKAPFSSSLKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGEP	720		
Db	661	ERGSVKYSAPETKAPFSSSLKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGEP	720		
Qy	721	STMGYWVALIILNSLCFLMNTIATYKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780		
Db	721	STMGYWVALIILNSLCFLMNTIATYKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780		
Qy	781	PVAFLSFSSLNLTFTISBEVIKFIILVVVPLPACINPLLYILFNPHKEDLSLRKQTYV	840		
Db	781	PVAFLSFSSLNLTFTISBEVIKFIILVVVPLPACINPLLYILFNPHKEDLSLRKQTYV	840		
Qy	841	WTRSKHPSLMSINSDVVKQSCDSTQALVFTTSSSITYDLPPSSVPSPAYPVTESCHLSS	900		
Db	841	WTRSKHPSLMSINSDVVKQSCDSTQALVFTTSSSITYDLPPSSVPSPAYPVTESCHLSS	900		
Qy	901	VAFVPCCL 907			
Db	901	VAFVPCCL 907			

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:30 ; Search time 104 Seconds
(without alignments)
4465.921 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSLGLVLLPLVLLQLAT.....PAYPTVTSCHLSVAFVPCFL 907

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4702	100.0	907	1	LGR5 HUMAN
2	4048	86.1	907	1	LGR5 MOUSE
3	2492.5	53.0	928	2	Q9BYD7
4	2412.5	51.3	923	2	Q86VU0
5	2410	51.3	915	2	Q6UY15
6	2091.5	44.5	951	1	LGR4 HUMAN
7	2088.5	44.4	951	1	LGR4 RAT
8	2024.5	43.1	927	2	Q8N537
9	1980	42.1	828	1	LGR6 HUMAN
10	1942.5	41.3	878	2	Q8BX59
11	1508	34.2	363	2	Q8C8A7
12	1182.5	25.1	1093	2	Q8HA06
13	1109.5	23.6	1257	2	Q7PNF8
14	1070	22.8	1360	2	Q7KTA0
15	1068	22.7	1360	2	Q8NDI1
16	1058.5	22.5	1012	2	Q95YI6
17	1058.5	22.5	1280	2	Q95YI7
18	976	20.8	1300	2	Q9NKD6
19	947.5	20.2	459	2	Q8R301
20	902.5	19.2	1050	2	Q9BN18
21	834.5	17.7	488	2	Q6PHA3
22	829.5	17.6	1039	2	Q86BL1
23	695.5	14.8	412	2	Q8BZR7
24	656.5	14.0	692	1	FSHR MOUSE
25	655	13.9	695	1	FSHR PIG
26	650	13.8	695	1	FSHR BOVIN
27	648.5	13.8	695	1	FSHR MACFA
28	642.5	13.7	694	1	FSHR HORSE
29	642	13.7	695	1	FSHR SHEEP
30	636	13.5	692	1	FSHR RAT
31	634.5	13.5	694	2	Q6YNB6

32	633	13.5	687	1	FSHR EQUAS
33	625.5	13.3	693	2	Q7ZTV5
34	625	13.3	688	2	Q64183
35	625	13.3	696	2	Q9DGF5
36	622	13.2	701	2	Q9DGC6
37	622	13.2	710	2	Q8AXM5
38	621.5	13.2	695	1	FSHR HUMAN
39	621	13.2	147	2	Q7IM96
40	616	13.1	694	2	Q6R6L8
41	616	13.1	695	2	Q8R428
42	606.5	12.9	704	2	Q6POA4
43	606.5	12.9	708	2	Q6TCF8
44	600	12.8	700	1	LSHR MOUSE
45	594	12.6	693	1	FSHR CHICK

ALIGNMENTS

RESULT 1

ID	LGR5 HUMAN	STANDARD;	PRT;	907 AA.
AC	Q75473; Q9UP75;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	Leucine-rich repeat-containing G protein-coupled receptor 5 precursor			
DE	(Orphan G protein-coupled receptor HG38) (G protein-coupled receptor			
DE	49)			
GN	Name=GPR49; Synonyms=GPR67, LGR5;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheraia; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98308104; PubMed=9642114; DOI=10.1006/bbrc.1998.8774;			
RA	McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T., Liu Q.;			
RT	"Identification and cloning of an orphan G protein-coupled receptor of			
RT	the glycoprotein hormone receptor subfamily."			
RL	Biochem. Biophys. Res. Commun. 247:266-270(1998).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Placenta;			
RX	MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;			
RA	Hsu S.Y., Liang S.-G., Hsueh A.J.W.;			
RT	"Characterization of two LGR genes homologous to gonadotropin and			
RT	thyrotropin receptors with extracellular leucine-rich repeats and a G			
RT	protein-coupled, seven-transmembrane region."			
RL	Mol. Endocrinol. 12:1830-1845(1998).			
CC	-!- FUNCTION: Orphan receptor. It may be an important receptor for			
CC	signals controlling growth and differentiation of specific			
CC	embryonic tissues (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal			
CC	cord, and various region of brain.			
CC	-!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.			
CC	-!- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
DR	EMBL; AF062006; AAC28019.1; -			
DR	EMBL; AF061444; AAC77911.1; -			
DR	PIR; J0176; J0176.			
DR	HSSP; Q9BZR6; IOZN.			
DR	GeneW; HGNC:4504; GPR49.			
DR	MIN; 606667; -			

DR GO: 0005887; C: integral to plasma membrane; TAS.
 DR GO: 0004930; F: G-protein coupled receptor activity; TAS.
 DR GO: 0007186; P: G-protein coupled receptor protein signalin. . .; TAS.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR InterPro: IPR002131; Gphrmn_receptor.
 DR InterPro: IPR001611; LRR.
 DR InterPro: IPR000372; LRR_Nterm.
 DR InterPro: IPR003591; LRR_typ.
 DR Pfam: PF00001; 7tm_1; 1.
 DR Pfam: PF00560; LRR; 16.
 DR Pfam: PF01462; LRRNT; 1.
 DR PRINTS: PRO0373; GLYCHORMONER.
 DR PRINTS: PRO0237; GPCR_Rhodopsin.
 DR PRINTS: PRO0019; LEURICHRPT.
 DR SMART: SM00369; LRR_TYP; 8.
 DR SMART: SM00013; LRRNT; 1.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1_1; FALSE_NEG.
 DR PROSITE: PS0262; G-PROTEIN RECEPTOR F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 907 Leucine-rich repeat-containing G protein-coupled receptor 5.
 FT DOMAIN 1 561 Extracellular (Potential).
 FT TRANSMEM 562 582 1 (Potential).
 FT DOMAIN 583 593 Cytoplasmic (Potential).
 FT TRANSMEM 594 614 2 (Potential).
 FT DOMAIN 615 638 Extracellular (Potential).
 FT TRANSMEM 639 659 3 (Potential).
 FT DOMAIN 660 682 Cytoplasmic (Potential).
 FT TRANSMEM 683 703 4 (Potential).
 FT DOMAIN 704 722 Extracellular (Potential).
 FT TRANSMEM 723 743 5 (Potential).
 FT DOMAIN 744 767 Cytoplasmic (Potential).
 FT TRANSMEM 768 788 6 (Potential).
 FT DOMAIN 789 802 Extracellular (Potential).
 FT TRANSMEM 803 823 7 (Potential).
 FT DOMAIN 824 907 Cytoplasmic (Potential).
 FT REPEAT 64 88 LRR 1.
 FT REPEAT 89 112 LRR 2.
 FT REPEAT 114 136 LRR 3.
 FT REPEAT 137 160 LRR 4.
 FT REPEAT 162 184 LRR 5.
 FT REPEAT 185 208 LRR 6.
 FT REPEAT 209 232 LRR 7.
 FT REPEAT 233 256 LRR 8.
 FT REPEAT 257 279 LRR 9.
 FT REPEAT 281 303 LRR 10.
 FT REPEAT 304 327 LRR 11.
 FT REPEAT 328 350 LRR 12.
 FT REPEAT 351 375 LRR 13.
 FT REPEAT 377 396 LRR 14.
 FT REPEAT 397 420 LRR 15.
 FT REPEAT 422 444 LRR 16.
 FT REPEAT 445 468 LRR 17.
 FT CARBOHYD 63 63 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 77 77 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 208 208 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 500 500 N-linked (GlcNAc. . .) (Potential).
 FT CARBOHYD 792 792 N-linked (GlcNAc. . .) (Potential).
 FT CONFLICT 90 90 R -> H (in Ref. 2).
 FT CONFLICT 212 212 L -> W (in Ref. 2).
 SQ SEQUENCE 907 AA; 99997 MW; 822D5C5E6F0D9092 CRC64;

Query Match 100.0%; Score 4702; DB 1; Length 907;
 Best Local Similarity 100.0%; Pred. No. 5.2e-272;
 Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLGVLVSLPVLLQATCGSSPRSGVLRGCPETHCHCEPDGRMLLRVDCSDLGISEL 60
 Db 1 MDTSLGVLVSLPVLLQATCGSSPRSGVLRGCPETHCHCEPDGRMLLRVDCSDLGISEL 60
 Qy 61 PSNLVSFTSYLDLSNMNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKLYM 120

Db 61 PSNLVSFTSYLDLSNMNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKLYM 120
 Qy 121 LQNNQLRHVPTEALQNLRLSLOSLRLDANHSIYVPPSCFSGLHSLRHLDNALTEIPVQ 180
 Db 121 LQNNQLRHVPTEALQNLRLSLOSLRLDANHSIYVPPSCFSGLHSLRHLDNALTEIPVQ 180
 Qy 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKCKCPGLHSLETLD 240
 Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKCKCPGLHSLETLD 240
 Qy 241 LNYNNLDEPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIDFVGRSA 300
 Db 241 LNYNNLDEPPTAIRTLNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPIDFVGRSA 300
 Qy 301 FOHLPELRTLTNGASQITFEFPLDTGTANLESILTGAQISSLSLQTCNQLPNLOVLDS 360
 Db 301 FOHLPELRTLTNGASQITFEFPLDTGTANLESILTGAQISSLSLQTCNQLPNLOVLDS 360
 Qy 361 YNLEDELPSFSVCOCKLOKIDLRHNEIYEIKVDTFQQLLSRLSLNLANWKIAIHPNAPST 420
 Db 361 YNLEDELPSFSVCOCKLOKIDLRHNEIYEIKVDTFQQLLSRLSLNLANWKIAIHPNAPST 420
 Qy 421 LPSLIKLDLSSNLSLSPFITGLHGLTHLKTGNHALQSLISSENFPBLKVIEMPYAYQCC 480
 Db 421 LPSLIKLDLSSNLSLSPFITGLHGLTHLKTGNHALQSLISSENFPBLKVIEMPYAYQCC 480
 Qy 481 AFGVCENAYKISQNWNGKDNSSMDLHKDAGMFOAQDERDLEDFLDLFEEDLKALHSVQ 540
 Db 481 AFGVCENAYKISQNWNGKDNSSMDLHKDAGMFOAQDERDLEDFLDLFEEDLKALHSVQ 540
 Qy 541 CSPSPGPFKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
 Db 541 CSPSPGPFKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
 Qy 601 AVNMLTGVSASVLAGVDAFTFGSPARHAGWENGVCVIGFLSIFASESVFLLTAAAL 660
 Db 601 AVNMLTGVSASVLAGVDAFTFGSPARHAGWENGVCVIGFLSIFASESVFLLTAAAL 660
 Qy 661 ERGSFVKYSAKFETKAPSSIKVILLICALLATMAAVPLLGSKYKASPLCLPLPGE 720
 Db 661 ERGSFVKYSAKFETKAPSSIKVILLICALLATMAAVPLLGSKYKASPLCLPLPGE 720
 Qy 721 STMGYVALILLNSLCLFLMWTIATKLYCNLDKDLNIWDCSMVKHIALILLFNCILNC 780
 Db 721 STMGYVALILLNSLCLFLMWTIATKLYCNLDKDLNIWDCSMVKHIALILLFNCILNC 780
 Qy 781 PVAFLSPSSILNLTFTISPEVIKFIILLVVVPLPACINPLLYILFNPFKEDLVSRKQTYV 840
 Db 781 PVAFLSPSSILNLTFTISPEVIKFIILLVVVPLPACINPLLYILFNPFKEDLVSRKQTYV 840
 Qy 841 WTRSKHPSLMSINSDVDVEKQCDSTQALVTFTSSITDLPSSVPSPAYVPTESCHLSS 900
 Db 841 WTRSKHPSLMSINSDVDVEKQCDSTQALVTFTSSITDLPSSVPSPAYVPTESCHLSS 900
 Qy 901 VAFVPCPL 907
 Db 901 VAFVPCPL 907

RESULT 2
 LGR5 MOUSE
 ID LGR5 MOUSE STANDARD; PRT; 907 AA.
 AC Q9Z1P4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
 DE (G protein-coupled receptor 49) (Orphan G protein-coupled receptor
 DE FEX).
 GN Name=Gpr49; Synonyms=Fex, Lgr5;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;


```

Qy 661 ERGFSVKYSAKPTKAPFSSKVLKIIICALLALTMAAVPLLGSKYKYGASPLCLPLPGE 720
Db 661 ERGFSVKYSAKPTKAPFSSKVLKIIICALLALTMAAVPLLGSKYKYGASPLCLPLPGE 720
Qy 721 STMGYMAVALIILNSLCFLMMTIAATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYMAVALIILNSLCFLMMTIAATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSFSNLNLFISBEVVKFILLVVPPLPACLNPLLYLFPNPKFEDLVSRKQTYV 840
Db 781 PVAFLSFSNLNLFISBEVVKFILLVVPPLPACLNPLLYLFPNPKFEDLVSRKQTYV 840
Qy 841 WTRSKHPSLMSINDVDEKSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Db 841 WTRSKHPSLMSINDVDEKSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Qy 901 VAFVPC 907
Db 901 VAFVPC 907

RESULT 3
Q9BYD7
ID Q9BYD7 PRELIMINARY; PRT; 928 AA.
AC Q9BYD7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE VTS20631 (Fragment).
GN Name=VTS20631;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki H., Hayaashi A., Kozuma S., Saito T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049405; BAB39854.1; -.
DR HSP; P25147; ID08.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gphrmu_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_type.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.
FT NON TER 1
SQ SEQUENCE 928 AA; 100487 MW; 4C3364ADEA89C463 CRC64;

Query Match 53.0%; Score 2492.5; DB 2; Length 928;
Best Local Similarity 55.7%; Pred. No. 3.4e-140;
Matches 491; Conservative 134; Mismatches 226; Indels 31; Gaps 9;

Qy 39 HCEPDGMLLRVDCSDGLSELPSNLVSFTSYDLSMNNISQLLPNPLPSLRFLEELRLA 98
Db 1 HCEPDG-IMLSADCSSELGSAPVGDLPDTAYLDLSMNNLTQLPGLFHFLFLEELRLS 59

Qy 99 GNALTYIPKGAFTGLYSKVLMLQNNQRLRHVPTAEALQNLRLQSLRLDANHSIVPPSCP 158
Db 60 GNHLSPICQAFSGLYSLKILMQNNQGLGPAEALWELPSLQSLRLDANLSLIPERSF 119

Qy 159 SGLSLRLHLDNALTIPVQAFPSLSAQMTIALANKIHIPYAFGNLSSLVVLHL 218
Db 120 EGLSLRLHLDNALTIPVQALNNLFPALQMTIALNKRISHIPYAFQNLTSVLVLHL 179

Qy 219 NNRHSLGKCFDGLHSLLETLDLNNYNNDEPPTAIRTLNKLGLGFHNNIRSIPEKAFV 278

```

```

Db 180 NNRHQLGTHSFEGHLNLTLDLNNYNLQBPFPVARTLGRLOELGFHNNNIKAIPKAFM 239
Qy 279 GNPSLIITHFYDNPIDQVGRSAFOHLPELRTLTINGASQITFPDPLGTANLESITLGA 338
Db 240 GNPLQTIHFYDNPIDQVGRSAFOYLPKLTHTLSUNGAMDIOEFPDLGTTSLLEITLTRA 299
Qy 339 QISSLPQTVCNQPNQVLDSLNNLEDPFSFVCQKLDKIDLRHNEIYEIKVDTFQOLL 398
Db 300 GIRLLPSGMCOQLPRLVLELSHNOIEELPSLHRCQKLEIGLQHNRIWEIGADTFSQLS 359
Qy 399 SLRSLNLAWNKIAIHPNAPSTLPSLKLDLSSNLLSPFTTGLHGTTHLKLQTNHALQS 458
Db 360 SLQALDLSWNAIRSIHPPEAFSTLSLVKLDLTNQLTTLPLAGLGLMHLKLGKLNLSQ 419
Qy 459 LLSSENPELVKVIEMPYAYOCCARGVCENAYKISNONKNGDNSSMDLH-----KKDA 511
Db 420 AFSKDSFPKRLILEVPYAYOCCPYGMCASFPKASQW-----BAEDLHLDDESSKRPL 473
Qy 512 GMFOAQDE---RDLEDFLLDDEEDLKALHSVQCSPPGPKCEHLIDGLWLRIGVWTI 567
Db 474 GLLARQAEHYDQDLDELQLEM-EDSKPHSPVQCSPTFGPKCEYLFEWGIRLAVWAI 532
Qy 568 AVLATCNALVTSTVFR-SPLYISPIKLLIGVIAAVNMLTGVSSAVLAGVDAFTFGSPAR 626
Db 533 VLLSVLCNGLVLLTVFAGGPAPLPVFKVVGAIAGANTLTGISGLLASVDALTFQFSE 592
Qy 627 HGAMWENGVCCHVIGFSLIFASESSVFLTLAALRGFSVKYSAKPTKAPFSSIKVILL 686
Db 593 YGARWETGLGCRATGFLAVLGSEASVILLTLAAVQCSVSVSCVRAYGKSPSLGSRVAGVL 652
Qy 687 LCALLALTMAAVPLLGSKYKYGASPLCLPL--PFGEPTMGYMAVALIILNSLCFLMMTIA 744
Db 653 GCLALAGLAALPLASVEYKASPLCLPYAPPEQPAALGFTVALVMNNSFCFLVAVAG 712
Qy 745 TKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNCPVAFISFSSLINLTIFISBEVVKFI 804
Db 713 IKLYCDLPRGDFEAVWDCAMVRHVAVLIFADGLLYCPVAFISFASMLGLFPVTPEAVKVS 772
Qy 805 LLVVPLPACLNPLLYLFPNPKFEDLVSRKQTYVWTRSKHPSLMSINDVDEKSCDSS 864
Db 773 LLVVPLPACLNPLLYLFPNPKFEDLVSRKQTYVWTRSKHPSLMSINDVDEKSCDSS 827
Qy 865 TQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSSVAFVPC 906
Db 828 TQALVAFSDVDLILEASEAGRP----FGLTYVGFPSVTLLSC 865

RESULT 4
Q86VU0
ID Q86VU0 PRELIMINARY; PRT; 923 AA.
AC Q86VU0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LGR6 protein (Fragment).
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood.
RX MEDLINE=22398257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

```


RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahay J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.L., Skalska U., Smallus D.E., Schmerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC047905; AAH47905.1; -;
DR HSP: O9BZ86; 1P8T
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016500; F: protein-hormone receptor activity; IEA.
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR002131; Gphrmn_receptor.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00560; LRR_1; 15.
DR PRINTS: PR00373; GLYCHORMONER.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00369; LRR_TYP; 8.
FT NON_TER 1
SQ SEQUENCE 923 AA; 100047 MW; 541D6746DAB06813 CRC64;
Query Match 51.3%; Score 2412.5; DB 2; Length 923;
Best Local Similarity 54.3%; Pred No. 2e-135;
Matches 483; Conservative 133; Mismatches 230; Indels 43; Gaps 10;
QY 32 RGCPTHCBCPDGRMLRLVDCSDGLSELPSNLVFTSYLDLSNMNISQLLPNPLPSLRF 91
DB 1 KGKPTG---QGMGRPLRLVLC-----QVSIISARDLSMNLTQLPGLPHLRF 47
QY 92 LEEELRAGNALTYPKGAFTGLYSKVLQNNQLRHVPTALQNLRLSRLDANHIS 151
DB 48 LEEELRAGNLTSHIPQAFSGLYSKILQNNQLGGIPAEALWELPSQLRLDANLIS 107
QY 152 YVPPSCFGLSLRLWLDNLTALTEIPVQAPFSLQALQNTLAKNIHHPDYAFGNLSS 211
DB 108 LVPERSFGLSLRLWLDNLTALTEIPVRLNLPALQALNTLAKNIHHPDYAFQNTLS 167
QY 212 LVVLHLHNNRHSGLKCKFDGLHSLTETLDLNNLDEFPPTAIRTLNLKELGFTSHNNIRS 271
DB 168 LVVLHLHNNRHSGLKCKFDGLHSLTETLDLNNLDEFPPTAIRTLNLKELGFTSHNNIRKA 227
QY 272 IPEKAFVGNPSLIITHFYDNPFIQVGRSAFOHLPELRLTLNGASQITEFPDPLTGANLE 331
DB 228 IPEKAFVGNPSLIITHFYDNPFIQVGRSAFOHLPELRLTLNGASQITEFPDPLTGANLE 287
QY 332 SILTGTGAISSIPQVCHQNLQVLDLSYNLLEPLSPSVCKQLKQIDLRHNEYIEIKV 391
DB 288 ILTLTRAGIRLLPSGMCOQLPRLRLVLESHNQIEELPSLRCKQKUEIGLQHNRIWEIGA 347
QY 392 DTFOQLLSRLSLNLANWKIATHNPAFTSLPSLKLDSNLSLSPFITGLHGLTHLKLIT 451
DB 348 DTFOQLLSRLSLNLANWKIATHNPAFTSLPSLKLDSNLSLSPFITGLHGLTHLKLIT 407
QY 452 GNHALQSLISSNPFELKVIEMPIYQCAFGVCENAYKINQWKNQNSMDDLH---- 507
DB 408 GNHALQSAFQKDSFPLRIEVPYAYQCCPYGMCASFQKASQW-----EAEIDLHLDDE 461
QY 508 ---KQDAGMFAQDE----RLEDFLLDFEEDLKALHSVQCSPPSPKPCPEHLLDGLWLI 560
DB 462 ESSKRPLGLLARAQENHYDQDLDELQLEM-EDSKPHPSVQCSPTPTGPKPCPEYLFESWGI 520

QY 561 RIGWTTIAVLATCNALVTSTVFR-SPLYISPIKLIKIGVIAVNMMLTGVSSAVLAGVDAF 619
DB 521 RLAVMAIYLLSGLNGLVLLTFVAGGPVLPFPVFGAIVAGANTLTGISCGLLASVDAL 580
QY 620 TFGSFARHAGWENGVCVHIGLSIPASESSVFLTLAALERGPSVKYSAKFETKAPFS 679
DB 581 TFGSFARHAGWENGVCVHIGLSIPASESSVFLTLAALERGPSVKYSAKFETKAPFS 640
QY 680 SLKVIILLCALLALMTAAVPLLLGGSKYASPLCLPL--PFGPESTMGVMVALILNLSLCP 737
DB 641 SVRAGVLGCLALAGLAAALPLASVGEYCASPLCLPYAPPEGQPAALGFTVALVMNSFCF 700
QY 738 LMWTIATKLYCNLDKGDLENINWDCSMVKHIALLLFTNCILNCPVAFSLFSSLIHLTFIS 797
DB 701 LVAGAYIKLYCDLPRGDFEAVWDCAMVRHVAWILFADGLLYCPVAFSLFASMLGLFPVT 760
QY 798 PEVIKFIILLVVVPLPACLNPLLYLLENPHFEDVLSLRKQTVWTRSKHPSLMSINSDV 857
DB 761 PEAVKSVLLVVLVPLPACLNPLLYLLENPHFEDVLSLRKQTVWTRSKHPSLMSINSDV 815
QY 858 EKQSCDSTQALVTFSTSSITYDLPPSSVPSPAYPVTVTESCHLSSVAFVPC 906
DB 816 EKQSCDSTQALVAFSDVDLILEASEAGRP----PGLETYGPSPVTLISC 860
RESULT 5
QYUY15
ID Q6UY15 PRELIMINARY; PRT; 915 AA.
AC Q6UY15
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Gonadotropin receptor.
GN ORFNames=UNQ6427;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
EX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heldens S.,
RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seahagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wiegand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment";
RL Genome Res. 13:2265-2270(2003).
DR EMBL: AY358119; AAQ88486.1; -;
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:0016500; F: protein-hormone receptor activity; IEA.
DR GO: GO:0004872; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR InterPro: IPR002131; Gphrmn_receptor.
DR InterPro: IPR001611; LRR.
DR InterPro: IPR003885; LRR_cyst.
DR InterPro: IPR003591; LRR_typ.
DR Pfam: PF00560; LRR_1; 15.
DR PRINTS: PR00373; GLYCHORMONER.
DR PRINTS: PR00237; GPCR_RHODOPSIN.
DR PRINTS: PR00019; LEURICHRPT.
DR SMART: SM00365; LRR_SD22; 5.
DR SMART: SM00369; LRR_TYP; 14.
KW Receptor.
SQ SEQUENCE 915 AA; 99265 MW; D57DD7A9DBB555F4 CRC64;


```
Query Match          51.3%; Score 2410; DB 2; Length 915;
Best Local Similarity 54.7%; Pred. No. 2.8e-135;
Matches 480; Conservative 130; Mismatches 227; Indels 40; Gaps 9;

Qy 44 GRMLLRVDCSLGISELPSNLVFTSYDLSMNNISQLLPNPLSLRPLELRAGNALT 103
Db 2 GRPLTLVC-----QVSIISARDLSMNNLTQLPGFLHRLFLBELRLSGNHL 51

Qy 104 YIPGATGTLGSLKVLMLONNQLRHVPTAEALONLSQSLRLDANHLISYPPSCFSLHS 163
Db 52 HIPQAFSGSLGSLKILMLONNQLGIPAEALWELPSQLSLRLDANHLISLVPERSFEGLS 111

Qy 164 LRHLWLDNALTETPVOAFRSLALQAMTLALNKHIPDYAFGNLSLVVLHNNRIH 223
Db 112 LRHLWLDNALTETPVRALNNLPAQAMTLALNKRISHIPDYAFQNLISLVVLHNNRIQ 171

Qy 224 SLGKCFDGLHSLRDLNLYNNLDEFFTAIRTLNKLKELGPHSNIRSIPEKAFVGNPSL 283
Db 172 HLGTHSFEGLHNLTLNLYNKLQEFFVAIRTLNKLQELGPHNNIRKAIPEKAFMGNPLL 231

Qy 284 ITHFYDNPTQVGRSAFOHLPELRTLTNGASQITTEFPDLTGTANLESILTLTCAQISL 343
Db 232 QTHIFYDNPTQVGRSAFOYLPKLTLSLNGAMDIQEPFDLKGTTSEILTLTRAGIRLL 291

Qy 344 PQTVCNLPNLQVLDLSNLELDPSPVCOCKLOKIDLRHNEIYEIKVDTPFOQLLSRSL 403
Db 292 PSGMCQQLPRURVLELHSHNQIEELPSLHRCQKLEIGLQHNRIWEIGADTFPSQLSSIQL 351

Qy 404 NLAWNKIAIHPNAFSTPLSLIKLDSNLSSPFTTGLHGLTHLKLGTGNHALQSLSSE 463
Db 352 DLSNNAIRSHPNAFSTPLSLIVSLKLDLTNDQTLTTLPLAGLGLMHLKGNLALSQAQSKD 411

Qy 464 NFPELKVEMPIYAOCCAFVCEYNAKISQNNKGDNSMDDLH-----KQDAGMFOA 516
Db 412 SFPKRLILEVYAYQCCPGYCMGASFCKASQGW-----EADLHLDDEESSKRLPLGLAR 465

Qy 517 QDE---RDLDEFLDDEEDLKALHSQCSPSPGPKCEHLLDGWLIRIGVWTIAVLAL 572
Db 466 QAENHYQDDLDELQLEM-EDSKPHSPVQCSPTPGFPKPCYLPESWGIRLAWAIVLLSV 524

Qy 573 TCNALVTSTVFR-SPLXIPIKLIIGVIAVNMMLTGVSASVAGVDAFTFGSPARHGAWW 631
Db 525 LCNGLVLLTVTPAGGPAPLPVKKVVGAGANTLTGICGLLASVDALTFCQFSEYGARW 584

Qy 632 ENGVCVHIGPLSTFASSESVFLTLAALRGFSVKYSAKETKAPSSSLKVIILLCALL 691
Db 585 ETGLGCRATGFLAVLGSEASVLLTLAAVQCSVSCVRAYKSPSLGSRAGVLGCLAL 644

Qy 692 ALTWAAVPLLGSKYKYGASPLCLPL--PRGEPSTWGYVMVALILLNSLCFLMMTIAVTKLYC 749
Db 645 AGLAAALPLASVGYGASPLCLCPVAPPEGQPAALGFTVALVMNMSFCFLVVGAGYIKLYC 704

Qy 750 NLDKGDLENIWDCSMVKHIALLLFTNCILNCFVAFSLFSSLINLTFISPEVIRKFIILVV 809
Db 705 DLPRGDFEAVWDCAMVRHVAWLIIPADGLLYCFVAFSLFASMLGLFPVTPPEAVKSVLLVVL 764

Qy 810 PLPACPLNLLYILNPHKEDLVSLRKQTYWTKSKHPSLMSINSDDEKQSCDSTQALV 869
Db 765 PLPACPLNLLYILNPHFRDRLRLRPRA-----GDSGPLAYAAAGLEKSSCDSTQALV 819

Qy 870 TFTSSSYDLPSSVSPSPAPVPTVESCHLSLVAFPVC 906
Db 820 AFSVDVLLILEASEAGRP-----PGLETYGPPSVTLISC 852

RESULT 6
LGR4 HUMAN STANDARD; PRT; 951 AA.
AC Q9EXB1; Q9NYD1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
```

```
DB (G protein-coupled receptor 48).
GN Name=GPR48; Synonyms=LGR4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=21294803; PubMed=11401528; DOI=10.1006/bbrc.2001.4625;
RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
RT "Molecular characterization of a novel glycoprotein hormone G-protein-
coupled receptor.";
RL Biochem. Biophys. Res. Commun. 282:757-764 (2001).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
placenta, ovary, testis and adrenal. Expressed also in spinal
cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate
and spleen.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch)
CC
CC EMBL: AF346711; AAK31153.1; -.
CC EMBL: AF346709; AAK31153.1; JOINED.
CC EMBL: AF346710; AAK31153.1; JOINED.
CC EMBL: AF257182; AAF68989.1; -.
CC HSP: Q9BZR6; LOZN.
CC GENE: HGNC:13299; GPR48.
CC MIM: 606666; -.
CC InterPro: IPR000276; GPCR_Rhodpsn.
CC InterPro: IPR002131; Gphrmn_receptor.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR000372; LRR_Nterm.
CC InterPro: IPR003591; LRR_typ.
CC Pfam: PF00001; 7tm.1; 1.
CC Pfam: PF00560; LRR_15.
CC Pfam: PF01462; LRRNT; 1.
CC PRINTS: PR00373; GLYCROMONER.
CC PRINTS: PR00237; GPCRHHODPSN.
CC PRINTS: PR00019; LEURICHRPT.
CC SMART: SM00013; LRRNT; 1.
CC SMART: SM00369; LRR_TYP; 4.
CC PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
CC PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
Signal; Transmembrane.
FT SIGNAL 1 24
FT CHAIN 25 951
FT Leucine-rich repeat-containing G protein-
coupled receptor 4.
FT DOMAIN 25 544
FT TRANSMEM 545 565
FT DOMAIN 566 575
FT TRANSMEM 576 596
FT DOMAIN 597 620
FT TRANSMEM 621 641
FT DOMAIN 642 661
FT TRANSMEM 662 682
FT DOMAIN 683 703
FT TRANSMEM 704 724
FT DOMAIN 725 756
FT TRANSMEM 757 777
FT DOMAIN 778 783
FT TRANSMEM 784 804
FT DOMAIN 805 951
```


FT REPEAT 55 79 LRR 1.
 FT REPEAT 81 103 LRR 2.
 FT REPEAT 104 127 LRR 3.
 FT REPEAT 128 151 LRR 4.
 FT REPEAT 153 175 LRR 5.
 FT REPEAT 176 199 LRR 6.
 FT REPEAT 201 223 LRR 7.
 FT REPEAT 224 247 LRR 8.
 FT REPEAT 248 270 LRR 9.
 FT REPEAT 272 294 LRR 10.
 FT REPEAT 318 341 LRR 11.
 FT REPEAT 342 366 LRR 12.
 FT REPEAT 368 387 LRR 13.
 FT REPEAT 388 411 LRR 14.
 FT REPEAT 413 435 LRR 15.
 FT DISULFID 618 693 By similarity.
 FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 294 294 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 505 505 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 292 292 F -> S (in Ref. 1; AAF68989).
 FT CONFLICT 433 433 L -> P (in Ref. 1; AAF68989).
 FT CONFLICT 668 668 L -> S (in Ref. 1; AAF68989).
 SQ SEQUENCE 951 AA; 104460 MW; 550C2DFC22CA1BB CRC64;

Query Match 44.5%; Score 2091.5; DB 1; Length 951;
 Best Local Similarity 50.7%; Pred. No. 2.9e-116;
 Matches 430; Conservative 124; Mismatches 271; Indels 23; Gaps 6;

QY 10 LSLPVLQLATGSSPRGVLRLGCPHCHCEPDGRLRLVDCSDGLSELPSNLVFTS 69
 DB 5 LGLLCLFALGLGSAGPSGAAPLCAAPCSDGD----RRVDCSGKGLTAVPEGLSAFTQ 60

QY 70 YLDLSMNNISQLPNPLSLRLEELRLAGNALTYIPKGAFTGLYSKVLMLQNNOLRHV 129
 DB 61 ALDISMNNITQPEDAFKNPPLEELQAGNDLSFIHPKALSGKLKELKVLTLQNNQLKTV 120

QY 130 PTEALQNLRSQSLDANHISYVPPSCFSGLSLRLHMLDNLALTEIPVQAFRSLALQ 189
 DB 121 PSEAIRGLSALQSLDANHITSVPEDSFEGVLQVLRHLWLDNLSLTPVPHPLSLPTLQ 180

QY 190 AMTLANKIHIPDYAFGNLSVLVHLNNRIHSLGKCPDGLHSLETLDLNNYNDLEF 249
 DB 181 ALTLANKISSIPDPAFTNLSSVLVHLNNKIRGLSQHCFDGLDNLLETLDLSYNNLGEF 240

QY 250 PTAITLNLKELGPHSNIRISPEKAFVGNPSLTTHFYDNPIDFVGRSAFQHLPELRT 309
 DB 241 PQAIRKARPSLKLGHSHNSISVIPDGFNDGNPLRLTHLYDNPISFVGNFAPNLSDLHS 300

QY 310 LTLNGASQITPEPDITGTANLESILTGAQISSLPQTVCNQLPNIQVLDLSYNLLEDLPS 369
 DB 301 LVIRGASVQQFPNLTGTVHLESILTGTAKISSIPNLCQEQKMLRLDLSYNNIRDLPS 360

QY 370 FSVCKLOKIDLRHNEIYEIKVDVTQQLLSRLSLANWKNKIAIHNPASFSTPLSLIKL 429
 DB 361 FNGCHALEEISLQRNIYQIKETGTFQGLSLRLDLSRLNLIHEIHSRAFATLGPITNLDV 420

QY 430 SSNLLSPFTGLHGLTHLKLGTGNHALQSLSSNPFELKVIEMPYAYOCCAFGCENAY 489
 DB 421 SFNETLSPFTTEGLNGLNKLVGNFKLKEALAKDFVNLRSLSVPYAYOCCAFWCDSYA 480

QY 490 KISQNWKNDSMDLHKQAGMFAQDQERDLEDFLDFEDLKALHS---VQCSPPG 546
 DB 481 NL----NTEDNSLDHSAVEKGTADAANVTS---TLNEEE----HSQIIHCTPTG 527

QY 547 PKPCEHLLDGMLIRIGVWTVIAVLTGNALVTSTVFRSPLYSIPKILIGVIAAVNMLT 606
 DB 528 AFKPCVYLLGSMWIRLTVFPLVALFFNLVILATFASCTSLPSSKLFGLISVSNLFM 587

QY 607 GVSSAVLAGVDAFTGSGFARHCAGWGVGCHVIGLFIASESSVFTLLTAAALRGFSV 666
 DB 588 GIYTGILTFLDVSWGFAEFGIWWETGSGCKVAGFLAVFSSESIFLLMLATVRSLSA 647

QY 667 KYSAKFETKAPFSSLSKVIILLCALLALTWAAVPLPGLGGSKYGLASPLCLPLPFGPSPMTGMY 726
 DB 648 KOIMKNGKSNHLKQPRVAALLAFLGATVAGCPFLPHRGYSASPLCLPPTGPTGTSLGFT 707

QY 727 VALIILNSICFLMWTIATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNCVPFLS 786
 DB 708 VTLVLLNSLAFLLMAVIYTKLYCNLEKEDLSNSQSMIKHVAWLIFTNCIFPCPVAFPS 767

QY 787 FSSILNLTPIGPEVIFKILLVVVPLPACLNPLLYLTFNPHFKEDLVSLRKQTYVWTRSKH 846
 DB 768 FAPLITATISPEIMKSVTLFFPLPACLNPLVYVFFNPKPKEDWKLKGRV---TKKSG 824

QY 847 PSLMSINS 854
 DB 825 SVSVSISS 832

RESULT 7
 LGR4 RAT
 ID LGR4 RAT STANDARD; PRT; 951 AA.
 AC Q922H4;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
 GN Name=Gpr48; Synonyms=Lgr4;
 OS Rattus norvegicus (Rat);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ovary;
 RA MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;
 RX Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
 RT "Characterization of two lgr genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region.";
 RL Mol. Endocrinol. 12:1830-1845(1998).
 CC -!- FUNCTION: Orphan receptor.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF061443; AAC77910.1; --
 DR HSSP; Q9BZR6; 10ZN.
 DR RGD; 628615; Gpr48.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR InterPro; IPR002131; Gprhm_receptor.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00001; 7tm_1; 1.
 DR Pfam; PF00560; LRR; 15.
 DR Pfam; PF01462; LRRNT; 1.
 DR PRINTS; PR00373; GLYCHORMONER.
 DR PRINTS; PR00237; GPCR_HODOPSIN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 5.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; FALSE_NEG.
 DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat; Signal; Transmembrane.

FT	SIGNAL	1	24	Potential.
FT	CHAIN	25	951	Leucine-rich repeat-containing G protein-coupled receptor 4.
FT				Extracellular (Potential).
FT	DOMAIN	25	544	1 (Potential).
FT	TRANSMEM	545	565	2 (Potential).
FT	DOMAIN	566	575	Cytoplasmic (Potential).
FT	TRANSMEM	576	596	2 (Potential).
FT	DOMAIN	597	619	Extracellular (Potential).
FT	TRANSMEM	620	640	3 (Potential).
FT	DOMAIN	641	661	Cytoplasmic (Potential).
FT	TRANSMEM	662	682	4 (Potential).
FT	DOMAIN	683	703	Extracellular (Potential).
FT	TRANSMEM	704	724	5 (Potential).
FT	DOMAIN	725	756	Cytoplasmic (Potential).
FT	TRANSMEM	757	777	6 (Potential).
FT	DOMAIN	778	783	Extracellular (Potential).
FT	TRANSMEM	784	804	7 (Potential).
FT	DOMAIN	805	951	Cytoplasmic (Potential).
FT	REPEAT	55	79	LRR 1.
FT	REPEAT	81	103	LRR 2.
FT	REPEAT	104	127	LRR 3.
FT	REPEAT	128	151	LRR 4.
FT	REPEAT	153	175	LRR 5.
FT	REPEAT	176	199	LRR 6.
FT	REPEAT	200	223	LRR 7.
FT	REPEAT	225	247	LRR 8.
FT	REPEAT	248	270	LRR 9.
FT	REPEAT	272	294	LRR 10.
FT	REPEAT	318	341	LRR 11.
FT	REPEAT	342	366	LRR 12.
FT	REPEAT	368	387	LRR 13.
FT	REPEAT	388	411	LRR 14.
FT	REPEAT	413	435	LRR 15.
FT	DISULFID	618	693	By similarity.
FT	CARBOHYD	68	68	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	188	188	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	199	199	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	294	294	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	314	314	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	505	505	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	951 AA;	104138 MW; EDD56AC072123461 CRC64;	
Query Match				
Best Local Similarity 44.4%; Score 2088.5; DB 1; Length 951;				
Matches 438; Conservative 118; Mismatches 271; Indels 19; Gaps 7;				
Qy	10	LSPVLQLATGSSPRGVLRLGCPHCHCEPDGRMLLRVDCSDLGSELPSNLSVPTS	69	
Db	5	LGLLCFLALGLSGAGSPGAAPPLCAAPCSGDG----	RRVDCSGKGLTAVPEGLSAFTQ	60
Qy	70	YLDLSMNNISQLLPNPLSLRFLBELRLAGNALTYIPKGAFTGLYSKLVLMQNNLRHV	129	
Db	61	ALDISMNNITQLPEDAFKSPFLEELQAGNDLSLIHPKALSGKELKVLTLQNNLRVT	120	
Qy	130	PTEALQNLRSQSLRLDANHISYPPSPSCFGLSHLRHLWLDNALTETPVQAFPSLSAQ	189	
Db	121	PSEAHGLSALQSLRLDANHISVPEDSFGLVQLRHLWLDNLSITEVVPPLSNLTLQ	180	
Qy	190	AMTLALNKIHHIPDYAFGNLSLVVLHLNNRIHSLGKKCFDGLHSLFETLDLNNLDEF	249	
Db	181	ALTALANNISIPDFAFNTLSVLVHLNNKIKLSQHCDFGLDNLLETLDLNNYLDEF	240	
Qy	250	PTAFTLSNLKELGFHNSNISIEKAFVGNPSLITHFYDNPITQFVCRSAFQHLPELRT	309	
Db	241	FOAIKALPSLKEGFHNSNISIVIPDAGFGNPLRLTHLYDNPISFVGNSAFNLSDLHC	300	
Qy	310	LTNLGASQITFPPDLTGTALESITLTGAQISSLPQTVCNQLPMLQVLDLSYNLLEDLPS	369	
Db	301	LVIRGASLVQFNNLTGTVHLESITLTGTKISSIPDDLCCQKQMLRLDLSYNNIRDLPS	360	
Qy	370	FSVCQKQKIDLRNHEIYEKVDTPQQLLSRLSNLANWKIAITHPNFSTLPSLIKLDL	429	
Db	361	FNGCRALBEISLQRNQSILKENTFQGLTSRLDLSRLNLSRIHSGAFKLTITNLDV	420	

Qy	430	SSNLLSFPITGLHGLTHLKLGTGNHALQSLISSNPFELKVIEMPYAYQCCAFGVCENAY	489	
Db	421	SFNEITSFPTEGLNGLNQLKLVGNFKLDALARDFANLRSLSVPIAYQCCAFGWCDSYA	480	
Qy	490	KISQNMKNKDNSSMDLHKKDAQMFQADRDLEDLFDLDFEEDLKALHSVQCSPPGPFK	549	
Db	481	NL----NTEDNSPQEHSVTKEG--ATDAANVTSTAENEHSQIIIH---CTPSTGAFK	530	
Qy	550	PCEHLDGWLIRGVWTVIAVLTALTCNALVTSTVPSPLIYSPKILGLGVIAAVNMLTGV	609	
Db	531	PCEYLLGSMWIRLTWVFIPLVALLFNLLVITVFASCSSLPASKLFIGLISVSNLLMGIY	590	
Qy	610	SAVLGAVDATTFGSFARHGAWGVCHVIGLISIFASBSVFLLLTALAERGFVSVKYS	669	
Db	591	TGILTFLDVAVSWGRFAEFGIWEYTGCKVAGSLAVFSSESAAVFLLLTAAVERSVFAKL	650	
Qy	670	AKFETKAPFSSLVKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGPESTGVMVAL	729	
Db	651	MKGKSSHLRQFQVAALLALLGAAGVCGFPLFHGGQYSASPLCLPFPPTGETPSLGTVTL	710	
Qy	730	ILNSLCFLAMTITATYTKLYCNLDKGLDLENWDCSMVGHIALLLFTNCILNCPVAFLSFS	789	
Db	711	VLLNSLAFLLMAIYITTKLYCNLEKEDLSNSQSSVVKHVAWLIFTCIFCPVAFPSFAP	770	
Qy	790	LINLTFTSPVPIKFIPLVPLPACLNPLLYILFNPFPKEDVLRLKQTVVWTRSKHPSL	849	
Db	771	LITRISISPIIMKSVTLIFPLPACLNPLVYVFNFPKEDKWLKKRV---TR-KHGSV	826	
Qy	850	-MSINS 854		
Db	827	SVSISS 832		
RESULT 8				
ID	Q8N537	PRELIMINARY;	PRT;	927 AA.
DT	01-OCT-2002	(TrEMBLrel. 22, Created)		
DT	01-OCT-2002	(TrEMBLrel. 22, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DE	GRP48	protein.		
OS	Homo sapiens	(Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,			
RA	Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,			
RA	Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,			
RA	Jones S.J., Marra M.A.			
RT	"Generation and initial analysis of more than 15,000 full-length human			
RT	and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			

Strausberg R.;
Submitted (JUN-2002) to the EMBL/GenBank/DDBJ databases.
EMBL; BC033039; AAH33039.1; -;
HSSP; P25147; 1DOB.
GO:0016021; C: integral to membrane; IEA.
GO:0016500; P: protein-hormone receptor activity; IEA.
GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
InterPro; IPR002076; GPCR_Rhodopsn.
InterPro; IPR002131; Gprhm_receptor.
InterPro; IPR001611; LRR.
InterPro; IPR000372; LRR_Nterm.
InterPro; IPR003591; LRR typ.
InterPro; IPR007087; Znf_C2H2.
Pfam; PF00001; 7tm_1; 1.
Pfam; PF01462; LRRNT; 1.
Pfam; PF00560; LRR_1; 14.
PRINTS; PR00373; GLYCHORMONER.
PRINTS; PR00237; GPCR_Rhodopsn.
PRINTS; PR00019; LEURICHRPT.
SMART; SM00013; LRRNT; 1.
SMART; SM00369; LRR_TYP; 5.
PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN 1.
SEQUENCE 927 AA; 101676 MW; C7B2F0C40E584C58 CRC64;
Query Match 43.1%; Score 2024.5; DB 2; Length 927;
Best Local Similarity 49.2%; Pred. No. 2.8e-112;
Matches 420; Conservative 125; Mismatches 254; Indels 55; Gaps 9;
6 LGVL--LSPVLLQLATGSSPRSGVLLRGCTHCHCEPDGEMLLRVDCSDLGLSELPSN 63
5 LGLLCFLAIGLGLGAGPGGAAPL-----CAAPCSGDG-----RRVDCSGKGLTAVPEG 54
64 LSVFTSYLDLSMNNTISQLLPNPLSLRPLEELRLAGNALTYIPKGAFTGLSLKVLMLQN 123
55 LSAT-----QALQAGNDLSFTHPKALSGELKVLTLQN 90
124 NQLRHVPTALQNLRSQSLRDANHSYVPPSCFSGHLSLRHLWLDNALTEIPVQAFR 183
91 NQLKTVPSAIRGLSALQSLRDANHSYVPPSCFSGHLSLRHLWLDNALTEIPVQAFR 150
184 SLSALQAMTLALNKHHPDYAFGNLSLAVLHNNRIHSLGKCPGLHSLTLDNY 243
151 NLPTLQALTLANKISSIPDFAFTNLSSLVHLHNNKIRSLSCFQGLDNLDTLDNY 210
244 NNLDEFPTAIRTLNKLKELGFHNSNIRSIPEKAFVGNPSLITIFPDNPIDFVGRSAFOH 303
211 NNLGEFFQAIKALPSLKLGFHNSISVIPDGFNGENPLTLIHYDNPISFVGNSAFHN 270
304 LPELRLTLNGASQITEPDLTGTLNLESLLTGAQISSLPQTVCNQLPNIQVLDLSYNL 363
271 LSDLSLVIRGASVMVQPPNLTGTVHLSLTGTGKISSIPNLCQEQKMLRTLDLSYN 330
364 LEDLPSVSCVQKQIDLRHNEIYEIKVDFTQQLLSRLSLMLANWKNIAIHPNFASTLPS 423
331 IRDLPSFNGCHALBEISLQRNIQVKEGTQGLISRLDLSRLNLIHEISRAFATLGP 390
424 LKLDLSNLSSPPTGLHGLTHLKTGNHALOSLISSENFPELKVTEMVYAYOCCAFG 483
391 ITNLDVSENLTSPTTEGLNGNLKLVGNFKLEALAKDFNLRSLSVPIYAYOCCAFW 450
484 VCENAYKISNOWNKDNMSMDLHKKDAGMFAQDERDLEDFLDLFBEDLKALHS---VQ 540
451 GCDSYANL---NTEDNSLDQSHVAQEGKTADAANVT---TLNEEB---HSQIIH 497
541 CSPSPGPKPEHLLDGLHIGVWTIAVLATCNALVTSTVPSPLYSIPKILGVIA 600
498 CTPSTGAFKPCYELGLSGMWIRLTWFIPLVALFNLLVITLTFASCTSLPSSEKLFGLIS 557
601 AVNMLTGVSVAVLGADVAFTRGSPARHCAGWENGVCVIGFLSTIFASESVFLLTLAAL 660
558 VSNLPMGIYTLTFLDVAWSGRFAEFGIWWETGSGCKVAGFLAVFSSESIFLLMLATV 617

QY 661 ERGFSVKYSAKFTYKAPFSSLKVILLCALLALTAWAVPLLGSGKYGASPLCLPLPFGEP 720
DB 618 ERLSIAKDIMKNGKSNHLKQFRVAALFLAGTAVAGCFPIFHRGYSASPLCLPFPPTGFT 677
QY 721 STMGYMVALIILNSLCPLMMTIAVTKLYCNLDKGLNWDSCVMKXHALLLFTNCILNC 780
DB 678 PSLGFTVTLVLSLAFLLMAVITTKLYCNLEKEDLSNSQSSMKHAWLLFTNCIPFC 737
QY 781 PVAFSLFSLINLTPISPEVIFKILLVVVPLACINLPLLYILFNPHFKEDLSLRKQTVY 840
DB 738 PVAFSPAPLITAISSPEIMKSVTLIFFPLACINLPLVYVFNPKFKEDKWLKLRV-- 795
QY 841 WTRSKHPSLMSINS 854
DB 796 -TKXSGSVSVSISS 808

RESULT 9
LGR6_HUMAN
ID LGR6_HUMAN STANDARD; PRT; 828 AA.
AC Q9HXB8; Q96KG9;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 6.
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20388592; PubMed=10935549; DOI=10.1210/me.14.8.1257;
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA van der Spek P.J., van Duin M., Heueh A.J.W.;
RT "The three subfamilies of leucine-rich repeat-containing G protein-
RT coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT signaling mechanism for LGR7.";
RL Mol. Endocrinol. 14:1257-1271(2000).
RN [2]
RP SEQUENCE OF 406-828 FROM N.A.
RX MEDLINE=22040266; PubMed=12044878; DOI=10.1016/S0014-5793(02)02775-8;
RA Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
RT "Identification of G protein-coupled receptor genes from the human
RT genome sequence.";
RL FEBS Lett. 520:97-101(2002).
RN [3]
RP SEQUENCE OF 431-828 FROM N.A.
RC TISSUE=Mammary gland;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Shiratori A., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yanazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hirakawa S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Horita T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Toguchi S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Taehiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y., Okamoto S.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,


```

Db      274  RRNQISEITSGALTNLRLKVLDDNSLSSMPVGLNMLLQEIASNNRIRVWSKGF 333
Qy      278  VGNESLTIHYDNPDIQVGRSAQHLPELRTLTNGASQITTEPDLTGTALESLLTG 337
Db      334  PKN--LVSLLDKSNPLAGIKFGALQNPRLKRLJLSDVRGLNELPPLDGCSTSEVRLDR 391
Qy      338  AQISLSPQVNCQNLQVLQVLDLSNLLLEDLPFSVCQKQKID----- 380
Db      392  ANLSKIDPHICKTSPRLSRLDKSNILLSINVTNCRDLRLDLASNRISLHGAPSSL 451
Qy      381  -----LRNHEIYEIKVDTFQQLSLRSRLNLANWKIAIHPNAPSTLPSPILKIDLSNL 433
Db      452  GQLHLLLSNNEIESIPHDAFVGLVRLQVLDMSNRVFFIHADAFPLKLEDLNGLNL 511
Qy      434  LSSPFTGLHGLTHLKTGNHALQSLSISENPELKVIEPMPYAVQCCA----- 481
Db      512  FPQLPTAGLELLHLKTFNNPHLREFPPEAPRIQTLVLSYAYHCCSFLLTALPKVP 571
Qy      482  --FGVCENA-YKISNQ-----WNKGDSN---SMDDLHKK-----DAGMFOA 516
Db      572  NTFIRENVLPPTNEPDMSLWNSYNDIWPQLQNLKSKFGTQINDLLNAYGAEVGSYPS 631
Qy      517  QDERLEDLDFEDLKAH-----SVQSPSPGPFKPCBHLLDGMLIRIGWNTIA 568
Db      632  GHVPTFPDEY--FEDELGITHASPTAOPGSIQCLPEPGFPLPCQDLFDWMTLRCGVVVF 689
Qy      569  VLALTCNALVTSTVFRSPLYISPIKLLIGVIAAANMLTGVSSAVLAGVDAFTGSPARHG 628
Db      690  LLAMLGNTVVVFLIFSRKMDVPRFLVCNLAARADFNGIYGLVAVVDASTLGEFRMYA 749
Qy      629  AWWENGVCVHIGLISFASBSSVFLTLTAAALRGFVKVSAKPEKAPSSSLKVIILLC 688
Db      750  IPWQMSAGCKSLGFLAVLSLSLVTYTLAVTLERNYALTHAMLNKRLSRHASVYMTVG 809
Qy      689  ALLALTAAPVLLGSKYKYGASPLCLPLPFG-PTMGVMVALIILNSLCFLMWTIATKL 747
Db      810  WTFATVAVLPLLVGSDYRVPAVCLPFEIQKGTGSLAYVVPFMFINGVAFILMGCYLKM 869
Qy      748  YCNLDKGDLENWDCSMVKHTALLFTNCILNCVAFILSPSLNLTAFISPEVIFKILLV 807
Db      870  YCAIRGSQAMNSDSRIAKRNWALLVFTDFICWSPIAFSLTAVFGLHLISLEQAKVFTVF 929
Qy      808  VVPLPACILNPLYILFNPHFKEDLVSLRKQ-----TVVTRSRKHPSLMSNSDDVEKQS 861
Db      930  ILPLNSCCNPLYALLTKQFKKQKCVLICKATEESRVTGIGCRHSNFS-----NRHT 983
Qy      862  CDSTOALVFTSSITVDLPP 882
Db      984  PANTNSLVERSCK---ELPP 1000

```

RESULT 14

Q7KTAO ID Q7KTAO PRELIMINARY; PRT; 1360 AA.

```

AC      Q7KTAO;
DT      05-JUL-2004 (TrEMBLrel. 27, Created)
DT      05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE      CG9830-PA (CG9830-pb).
OS      Drosophila melanogaster (Fruit fly).
OC      Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC      Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OK      NCBI_taxID=7227;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA      Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.E.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA      Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,

```

```

RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA      Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borkova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,
RA      Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA      Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA      Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA      Hostali D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA      Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA      Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA      Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA      Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA      Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA      Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhou X., Zhu S.,
RA      Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT      "The genome sequence of Drosophila melanogaster.";
RL      Science 287:2185-2195 (2000).
RN      [2]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22426065; PubMed=12537568;
RA      Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA      Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA      George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
RA      Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA      Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA      Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT      "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT      melanogaster euchromatic genome sequence.";
RL      Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22426070; PubMed=12537573;
RA      Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA      Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA      Ashburner M., Celniker S.E.;
RT      "The transposable elements of the Drosophila melanogaster euchromatin:
RT      a genome perspective.";
RL      Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN      [4]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=22426069; PubMed=12537572;
RA      Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA      Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochman S.E.,
RA      Smith C.D., Tupy J.L., Whittier E.J., Bayraktaroglu L., Berman B.P.,
RA      Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA      Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA      Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA      Lewis S.E.;
RT      "Annotation of the Drosophila melanogaster euchromatic genome: a
RT      systematic review.";
RL      Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN      [5]
RP      SEQUENCE FROM N.A.
RX      FlyBase;
RL      Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN      [6]
RP      SEQUENCE FROM N.A.

```



```
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB003642; AAF53367.3; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016500; F: protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gprhrn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003885; LRR_cyst.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 1.
DR Pfam; PF00560; LRR_1; 16.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00365; LRR_SD22; 7.
DR SMART; SM00369; LRR_TYP; 14.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 1360 AA; 150772 MW; 5B5BC47C933CD20 CRC64;

Query Match 22.8%; Score 1070; DB 2; Length 1360;
Best Local Similarity 28.9%; Pred. No. 3.9e-55;
Matches 282; Conservative 153; Mismatches 423; Indels 118; Gaps 14;

QY 17 QLAT-----GGSSPRSGVLLRGCPHCHCEPDGRM-LLRVDCSDGLGSELPSNLSPV 67
DB 123 QLSTASTATDVIAGKRTLHSICERWLQKHC--TGSLEVLRLSCRGIGILAVPVLNPN 180
QY 68 TSYLDLNNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKLKVLMLQNNQIR 127
DB 181 VVVDLGNNTITKLEANSFFMAPNLELTSDNSINMDPNFYGLAKRLSLQNCGLK 240
QY 128 HVPTEALQNLRLDANHSIVPPSCFSLHSLRLHMDNLTPEIPVQAFRLSA 187
DB 241 SLPPSQFGLAQLTSLQNGNALVSLDGDCLGHLQKLTLEGLNLFRIPTNALAGLRT 300
QY 188 LOAMTLALNKHIPDYAFGNLSSIVLHLHNNRHISLGKCFDGLHSLTLDLNNYLD 247
DB 301 LEALNLSGSLTINDEDFRPNLIVLLKRNQIMKISAGALKNLTKALKVLELDNLTIS 360
QY 248 EPTAIRTSLNKLGEFFSNIRSIPEKAFVGNPSLIITHFYDNPIDQFVGSAPHLPEL 307
DB 361 SLPEGLSKLSQQLSLSITSNRLINDTELPRSMQMDMR--ANPLSTISPGAFRMSKL 418
QY 308 RLTINGASQITFPDLGTANLESLTLTGQISSLPTQCNQLPNQLVLDLSYNLLEDL 367
DB 419 KKLILSDVTRLSRPFLEACHALEILKLDRAQIEVPANLCRQTPRLKSLKLTNSLKRI 478
QY 368 PSFVSCQKQKIDLRHNIYEIK-----VDTFQQLSLRSL 403
DB 479 PNLSSCDRLRLDLSNNQIEKIQQKPPNGLKQLNDLLSYNRIKALPQDAFQGIPLQLL 538
QY 404 NLANKIAIHPNAPSTPLSLKLDLSNLSLSPFITGLHGLTHLTKGNHALQSLISSE 463
DB 539 DLEGEISYIHKFAFSGFTALDNLGNINIPPELPSGLRALHLKTFNPNKLRFPFPPD 598
QY 464 NPPELKVTEMPAYOCCAF-----GVCE 486
DB 599 TFPRIQTILSYAHCCAFLEPLVAMSSQKTSQVQEAFLVPSDAEFMTLNNSMMNIWP 658
QY 487 NAYKISNQ-----WNKGDNSMDDLHKKDKAG-----MFQADQERDLEDLDF--- 529
DB 659 QMHNLSKQLGASMDHPWETAINFNEEQLOTQTGGQIATSYMEYFEEDHVSQATGYGFG 718
QY 530 -----EEDLKALHSVQCSFSPGPPPCHELLDGLWIRIGWTVIAVLATTCNALVTST 581
DB 719 TGLFSGMSTEDFQP--GSVQCLFMPGPFPLFCADLFDWMTLRGCVWVVFLLSLNGTVVVF 777
QY 582 VFRSPLYISPIKLLIGVTAANVNLTVGVSSAVLAGVDAAFTFGSPARHGAWENGVCCHVIG 641
DB 778 LLCRSRKMVPRFLVCLNLAADFFPMGIYLGILAIVDAATLGEFRFAIPQMSVLCQUSG 837
```

```
QY 642 FLSIPASSSVFLLTLAALRGFSVKYSAKETKAPFSSLKVIILLCALLATMAVPLL 701
DB 838 FLAVLSSELSVYTLAVITLERNYAIATHLNLKRLSLKQAGYIMGVWVFALIMALPLV 897
QY 702 GSKYSGASPLCLPLPFGG-PSTMGYMVALILNLSLCELMMTIATYTKLYCNLDKGLNITW 760
DB 898 GVSDYRKFAVCLPFPETTTGPASLTIVISLMFINGCAFLTLMGCYUKMTWAIKGSQAWNTN 957
QY 761 DCSMVKHIALLFTNCILNCIPVAFLSFSLINLTFISPEVINKFILLVVVPLPACLNPLLY 820
DB 958 DSRIAKRMALLVFTDFLCWSPIAFPSITAFGLQLISLEQAKIFTVFVLPLNSCCNPPLY 1017
QY 821 ILFNPHFKEDLVSLRKQTYVWTRSKHPSLMSINSDDVEKQSCDSQALYTFSSSITYDL 880
DB 1018 AIMTKQPKKDCVTL-----CKH-----FEESRVVGGGPGGRGAVARTKRG---DL 1060
QY 881 PPSSVPSPAYPVWTESC 896
DB 1061 PPPLLPAAAAVAHPGCG 1076

RESULT 15
QSNDDII PRELIMINARY; PRT; 1360 AA.
AC Q9NDII;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Glycoprotein hormone receptor II.
GN Name=rk;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=Canton S.; TISSUE=Whole animal;
RX MEDLINE=20359836; PubMed=10899142; DOI=10.1101/gr.10.7.924;
RA Eriksen K.K., Hauser F., Schiott M., Pedersen K.-M., Soendergaard L.,
RA Grimmelikhuijzen C.J.P.;
RT "Molecular cloning, genomic organization, developmental regulation,
RT and a knock-out mutant of a novel leu-rich repeats-containing G
RT protein-coupled receptor (DLGR-2) from Drosophila melanogaster.";
RL Genome Res. 10:924-938(2000).
DR EMBL; AF142343; AAF66608.1; -.
DR HSSP; Q9B2E6; 1OZN.
DR FlyBase; FBgn0003255; rk.
DR GO; GO:0007564; P: regulation of cuticle tanning; IMP.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gprhrn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00001; 7cm_1; 1.
DR Pfam; PF00560; LRR_1; 16.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 1360 AA; 150732 MW; 7D435155B4F6F612 CRC64;

Query Match 22.7%; Score 1068; DB 2; Length 1360;
Best Local Similarity 28.8%; Pred. No. 5.1e-55;
Matches 281; Conservative 154; Mismatches 423; Indels 118; Gaps 14;

QY 17 QLAT-----GGSSPRSGVLLRGCPHCHCEPDGRM-LLRVDCSDGLGSELPSNLSPV 67
DB 123 QLSTASTATDVIAGKRTLHSICERWLQKHC--TGSLEVLRLSCRGIGILAVPVLNPN 180
QY 68 TSYLDLNNISQLLPNPLPSLRFLELRAGNALTYIPKGAFTGLYSKLKVLMLQNNQIR 127
```


Db 181 VVVDLGNNNLTKEANSFFMAPNLEDLTLSDNSIINMDPNAFYGLAKLKELSLQNCGLK 240

Qy 128 HVPTEALQNRSLQSLRLDANHIYVPPSCFSLHSLRHLWLDNALTEIPVQAFRSLSA 187

Db 241 SLPPQSQGLAQLTSLQINGNALVSLDGDCLGHQKRLTLRLEGNLFYRIPTNALAGLRT 300

Qy 188 LOAMTLALNKHHTIPDYAFGNLSLVLVHLHNNRHSIGKCKPDGLHSLFETLDLNYNNLD 247

Db 301 LEALNLSNLITIINDEDFPPMPNLI VLLUKRNQIMKISAGALKNLTKALVLELDDNLIS 360

Qy 248 EFTAIRTLNKLGFHSHNIRSIPEKAFVGNPSLITIIHFDNPIQFVGSRSAFOHLPEL 307

Db 361 SLPEGLSKLSQQLSITSNLRWINDELPMSQMLDMR--ANPLSTISAGAFRGMNKL 418

Qy 308 RLTTLNGASQITEPPDLTGTANLESFLTGAQISSLPQTVCNQLPNIQVLDLSYNLLEDL 367

Db 419 RKJLTSVVRTLRSPELEACHALEILKDRAGIQEVPANLCROTTPRLKSLKELKTNLSKRI 478

Qy 368 PSESVCQKQKIDLRHNEIYEIK-----VDTFQQLLSLSL 403

Db 479 PNLSSCRDLRLDLSNQIEKIQKPPNGLKQLNDLLSNRIKALPQDAPQGIKQLLL 538

Qy 404 NLAWNKIAIHPNAFSLPSLIKLDSNLSSPFIITGLHGLTHLKLTHGHALQSLISSE 463

Db 539 DLEGNEISYIHKEAFSGFTALEDNLGNIPPELPESGLRALHLKTFNPKLREFPPPD 598

Qy 464 NPPELKVIEPYYQCCAF-----GVCE 486

Db 599 TFPRIQTLILSYAHCCAFPLVAMSSQKTSQVQEAFLPSPDAEFDMTLMNNSMMNIWP 658

Qy 487 NAYKISNQ-----NKGDNSSMDLHKKQAG-----MFOAQDERDLEDLDF--- 529

Db 659 QMHNLSKQLGASMDPNETAINFNEEQLOQTGQIATSYNEEYFEEHDSGGPATGYGFG 718

Qy 530 -----EEDLKALHSVQSPSPGPKCEHLLDGLIRIGVWTIAVLALTCNALVTST 581

Db 719 TGLPSGMSTEDFQPGSVQCLPMFGPLPCADLPDMWTLRCGVWVFLSLGNGTVFV 777

Qy 582 VFRSPLYISPIKLLIGVIAVNMUTGVSSAVLAGVDAFTPGSPARHGAWENGVGCHVIG 641

Db 778 LLCRSKMDVPRFLVNCNLAADFFWGIYLGILAIVDAATLGEFRMFAIPWQMSVLCQLSG 837

Qy 642 FLSTIFASESSVFLTLAALERGPSVKYSAKFETKAPFSSLSKVIILLCALLALTMAAVPLL 701

Db 838 FLAVLSSELSVYTLAVITLERNYAITHAHUNKLSLKQAGYINSGVWFALINMLPLV 897

Qy 702 GGSYKYGASPLCLPFGGE-PSTMGYVVALIILNSLCFLMMTIAYTKLYCNLDKGDLENIW 760

Db 898 GVS DYKFAVCLPFTTTGPASTYVISLMPINGCAFLTLMGCYLKMYWAIRGSQAWNTN 957

Qy 761 DCSMVKHIALLLFTNCILNCVPVAFLSFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLY 820

Db 958 DSRIAKRMALIVFTDFLCWSPIAFESITAFGLQISLEQAKIFTVFVPLNSCCNPFLY 1017

Qy 821 ILFNPHEKDLVSLRKQTYVWTRSKHPSLMSINSDDEKQSCDSTQALVTFSSSITYDL 880

Db 1018 AINTKQFKKOCVIL-----CKH-----FEESRVVGGGPGCGRAVARTKKG---DL 1060

Qy 881 PPSSVPSPAYPVTESC 896

Db 1061 PPPLPAAVAHPPGC 1076

Search completed: July 12, 2005, 07:36:34
Job time : 115 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:32 ; Search time 614 Seconds
(without alignments)
570.689 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSELGLVLLSLPVLQLAT.....PAYPTESCHLSLVAFPVCL 907

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1726218 seqs, 386331768 residues

Total number of hits satisfying chosen parameters: 1726218

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PTC_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PTCUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US11A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4702	100.0	907	10	US-09-965-536A-16
2	4702	100.0	907	10	US-09-970-944-26
3	4702	100.0	907	14	US-10-251-385-264
4	4702	100.0	907	14	US-10-225-567A-422
5	4702	100.0	907	15	US-10-295-027-849
6	4702	100.0	907	15	US-10-295-027-946
7	4702	100.0	907	15	US-10-295-027-1331
8	4702	100.0	907	16	US-10-751-736-84
9	4702	100.0	907	17	US-10-482-029-158
10	4701	100.0	1145	17	US-10-505-486-103
11	4696	99.9	907	14	US-10-251-385-278

Query Match 100.0% Score 4702; DB 10; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

12	4591	99.8	907	10	US-09-970-944-27	Sequence 27, Appl
13	4586	96.9	883	15	US-10-295-027-484	Sequence 484, App
14	4586	96.9	883	15	US-10-295-027-1330	Sequence 1330, Ap
15	4586	96.9	883	15	US-10-173-999-28	Sequence 28, Appl
16	4048	86.1	907	14	US-10-271-078-10	Sequence 10, Appl
17	3598	76.5	693	17	US-10-851-470-4	Sequence 4, Appli
18	2534	53.9	967	10	US-09-851-595-2	Sequence 2, Appli
19	2534	53.9	967	15	US-10-684-667-2	Sequence 2, Appli
20	2527.5	53.8	1205	17	US-10-505-486-64	Sequence 64, Appl
21	2525.5	53.7	967	10	US-09-851-595-11	Sequence 11, Appl
22	2525.5	53.7	967	15	US-10-331-496A-95	Sequence 95, Appl
23	2525.5	53.7	967	15	US-10-664-667-11	Sequence 11, Appl
24	2525.5	53.7	967	16	US-10-737-450-32	Sequence 32, Appl
25	2525.5	53.7	1005	14	US-10-176-847-90	Sequence 90, Appl
26	2492.5	53.0	928	10	US-09-970-944-23	Sequence 23, Appl
27	2465	52.4	940	10	US-09-970-944-6	Sequence 6, Appli
28	2413	51.3	915	15	US-10-295-027-1336	Sequence 1336, Ap
29	2413	51.3	915	16	US-10-783-528-101	Sequence 101, App
30	2410	51.3	915	15	US-10-331-496A-68	Sequence 68, Appl
31	2409	51.2	915	14	US-10-270-336-5	Sequence 5, Appli
32	2389	50.8	893	10	US-09-970-944-34	Sequence 24, Appl
33	2246	47.8	948	16	US-10-398-036-10	Sequence 10, Appl
34	2106.5	44.7	951	17	US-10-935-190-5	Sequence 5, Appli
35	2103.5	44.7	1189	17	US-10-505-486-126	Sequence 126, App
36	2094.5	44.5	951	17	US-10-851-470-2	Sequence 2, Appli
37	2088.5	44.0	951	14	US-10-276-340-3	Sequence 3, Appli
38	2070.5	44.0	951	14	US-10-225-567A-528	Sequence 528, App
39	1980	42.1	823	14	US-10-270-336-6	Sequence 6, Appli
40	1980	42.1	828	14	US-09-970-944-25	Sequence 25, Appl
41	1980	42.1	828	14	US-10-225-567A-581	Sequence 581, App
42	1374	42.0	794	14	US-10-270-336-7	Sequence 7, Appli
43	1373	42.0	828	14	US-10-270-336-2	Sequence 2, Appli
44	1844	39.2	736	10	US-09-851-595-8	Sequence 8, Appli
45	1844	39.2	736	15	US-10-664-667-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-965-536A-16
; Sequence 16, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J. N.
; APPLICANT: MINTIER, G.
; APPLICANT: RAMANATHAN, C. S.
; APPLICANT: HAWKEN, D. R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY5,
; FILE REFERENCE: D0041NP
; CURRENT APPLICATION NUMBER: US/09/965,536A
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,713
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/261,781
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/306,605
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 60/310,436
; PRIOR FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 16
; LENGTH: 907
; TYPE: PRT
; ORGANISM: HUMAN
US-09-965-536A-16


```
QY 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGISEL 60
Db 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGISEL 60
QY 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRFLEBLRLAGNALTYIPKGAFTGLYSLKVLM 120
Db 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRFLEBLRLAGNALTYIPKGAFTGLYSLKVLM 120
QY 121 LQNNQLRHVPTEALQNLSLQSLRLDANHI SYVPPSCFSGLSLHRLWLDNNALEIPIVQ 180
Db 121 LQNNQLRHVPTEALQNLSLQSLRLDANHI SYVPPSCFSGLSLHRLWLDNNALEIPIVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRHSIGLKKCFDGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRHSIGLKKCFDGLHSLETL 240
QY 241 LNYNNLDEFPPTAIRTLNLSKELGFHNSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
Db 241 LNYNNLDEFPPTAIRTLNLSKELGFHNSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
QY 301 FOHLPELRTLINGASQITEPPDLTGTANLESLETLTGAQISSLPQTVNCQNPQLQVLDLS 360
Db 301 FOHLPELRTLINGASQITEPPDLTGTANLESLETLTGAQISSLPQTVNCQNPQLQVLDLS 360
QY 361 YNLLEDLPFSVCOKLOKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIAI IHPNAPST 420
Db 361 YNLLEDLPFSVCOKLOKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIAI IHPNAPST 420
QY 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKL TGNHALQSLI SSENPELKIEMPVAYOCC 480
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKL TGNHALQSLI SSENPELKIEMPVAYOCC 480
QY 481 AFGVCENAYKISNQWNGKDNSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISNQWNGKDNSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCBHLDDGWLIRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGWLIRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWWENGVCVHIFGLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWWENGVCVHIFGLSIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSAKFETKAPSSSLKVIILLCALLATWAAVPLLGSGSKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPSSSLKVIILLCALLATWAAVPLLGSGSKYGASPLCLPLPFGEP 720
QY 721 STMGYMWALILNLSLCLFMMTIAVTKLYCNLDKGDLNIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYMWALILNLSLCLFMMTIAVTKLYCNLDKGDLNIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFISPEVIKFI LLVVVPLPACLNPLLYILFNPHFKEDVLSLRKQTYV 840
Db 781 PVAFLSFSSLINLTFISPEVIKFI LLVVVPLPACLNPLLYILFNPHFKEDVLSLRKQTYV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSSITVDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPC 907
Db 901 VAFVPC 907
```

RESULT 2

US-09-970-944-26
; Sequence 26, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shinkets, Richard A

; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 26
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-26

Query Match 100.0%; Score 4702; DB 10; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGISEL 60
Db 1 MDTSRGLVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGISEL 60
QY 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRFLEBLRLAGNALTYIPKGAFTGLYSLKVLM 120
Db 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRFLEBLRLAGNALTYIPKGAFTGLYSLKVLM 120
QY 121 LQNNQLRHVPTEALQNLSLQSLRLDANHI SYVPPSCFSGLSLHRLWLDNNALEIPIVQ 180
Db 121 LQNNQLRHVPTEALQNLSLQSLRLDANHI SYVPPSCFSGLSLHRLWLDNNALEIPIVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRHSIGLKKCFDGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLVHLHNNRHSIGLKKCFDGLHSLETL 240
QY 241 LNYNNLDEFPPTAIRTLNLSKELGFHNSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
Db 241 LNYNNLDEFPPTAIRTLNLSKELGFHNSNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
QY 301 FOHLPELRTLINGASQITEPPDLTGTANLESLETLTGAQISSLPQTVNCQNPQLQVLDLS 360
Db 301 FOHLPELRTLINGASQITEPPDLTGTANLESLETLTGAQISSLPQTVNCQNPQLQVLDLS 360
QY 361 YNLLEDLPFSVCOKLOKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIAI IHPNAPST 420
Db 361 YNLLEDLPFSVCOKLOKIDLRHNEIYEIKVDYTFQQLSLSLNLANWKIAI IHPNAPST 420
QY 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKL TGNHALQSLI SSENPELKIEMPVAYOCC 480
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKL TGNHALQSLI SSENPELKIEMPVAYOCC 480
QY 481 AFGVCENAYKISNQWNGKDNSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISNQWNGKDNSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCBHLDDGWLIRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGWLIRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWWENGVCVHIFGLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWWENGVCVHIFGLSIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSAKFETKAPSSSLKVIILLCALLATWAAVPLLGSGSKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPSSSLKVIILLCALLATWAAVPLLGSGSKYGASPLCLPLPFGEP 720
QY 721 STMGYMWALILNLSLCLFMMTIAVTKLYCNLDKGDLNIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYMWALILNLSLCLFMMTIAVTKLYCNLDKGDLNIWDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFISPEVIKFI LLVVVPLPACLNPLLYILFNPHFKEDVLSLRKQTYV 840
Db 781 PVAFLSFSSLINLTFISPEVIKFI LLVVVPLPACLNPLLYILFNPHFKEDVLSLRKQTYV 840

Db 781 PVAFSLFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYLILFNPHFKEDLVSLRKQTYV 840

Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITTYDLPPSSVPSPAYPVTESCHLSS 900

Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITTYDLPPSSVPSPAYPVTESCHLSS 900

Qy 901 VAFVPCPL 907

Db 901 VAFVPCPL 907

RESULT 3

US-10-251-385-264

; Sequence 264, Application US/10251385

; Publication No. US20030105292A1

; GENERAL INFORMATION:

; APPLICANT: Behan, Dominic P.

; APPLICANT: Chalmers, Derek T.

; APPLICANT: Liaw, Chen W.

; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G

; TITLE OF INVENTION: Protein-Coupled

; TITLE OF INVENTION: Receptors

; FILE REFERENCE: AREN-0040

; CURRENT APPLICATION NUMBER: US/10/251,385

; PRIOR FILING DATE: 2002-09-20

; PRIOR FILING DATE: 1998-10-13

; NUMBER OF SEQ ID NOS: 294

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 264

; LENGTH: 907

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-251-385-264

Query Match 100.0%; Score 4702; DB 14; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSRGLVLLSLPVLQLATGSSPRSGVLLRGCPCHCEPDGGRMLLRVDCSDLGSEL 60

Db 1 MDTSRGLVLLSLPVLQLATGSSPRSGVLLRGCPCHCEPDGGRMLLRVDCSDLGSEL 60

Qy 61 PSNLVSFTSYDLDSNNNISQLLPNPLPSLRFLERLGNALTYIPKGAFTGLYSLKVL 120

Db 61 PSNLVSFTSYDLDSNNNISQLLPNPLPSLRFLERLGNALTYIPKGAFTGLYSLKVL 120

Qy 121 LQNNQLRHVPTEALQNLRLDANHI SYVPPSCFSGLSRLHRLWLDNLTETIPVQ 180

Db 121 LQNNQLRHVPTEALQNLRLDANHI SYVPPSCFSGLSRLHRLWLDNLTETIPVQ 180

Qy 181 AFRSLSALOQMTALNKTHIIPDYAFGNLSLVHLHNNRIHSLGKCFDGLHSLTLD 240

Db 181 AFRSLSALOQMTALNKTHIIPDYAFGNLSLVHLHNNRIHSLGKCFDGLHSLTLD 240

Qy 241 LNNYNDLDFPTAIRTLNKLGPNSNNIRSIPEKAFVGNPSLITIFDYNPIQVGRSA 300

Db 241 LNNYNDLDFPTAIRTLNKLGPNSNNIRSIPEKAFVGNPSLITIFDYNPIQVGRSA 300

Qy 301 FOHLPELRTLNGASQITTEPDLTGNTANLESILTGAQISLSPQTCVQNLQVLDLS 360

Db 301 FOHLPELRTLNGASQITTEPDLTGNTANLESILTGAQISLSPQTCVQNLQVLDLS 360

Qy 361 YNLLEDLPFSVCOKLQIDLRHNEIYEIKVDYTFQOLLSLSLNLANWKIAIHPNAPST 420

Db 361 YNLLEDLPFSVCOKLQIDLRHNEIYEIKVDYTFQOLLSLSLNLANWKIAIHPNAPST 420

Qy 421 LPSLTKLDLSSNLSFFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPVAQCC 480

Db 421 LPSLTKLDLSSNLSFFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPVAQCC 480

Qy 481 AFGVCENAYKISNOWNKGDNSMDDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540

Db 481 AFGVCENAYKISNOWNKGDNSMDDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540

Qy 541 CSPSPGPKPCEHLLDGLIRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Db 541 CSPSPGPKPCEHLLDGLIRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Qy 601 AVNMLTGUSSAVLAGVDATFGSFAHCAWENGVCCHVIGFLSIFASESSVFLTLTAAAL 660

Db 601 AVNMLTGUSSAVLAGVDATFGSFAHCAWENGVCCHVIGFLSIFASESSVFLTLTAAAL 660

Qy 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALLALTAAPVPLLGSGKYGASPLCLPFGEP 720

Db 661 ERGFSVKYSAKFETKAPFSSLSKVIILLCALLALTAAPVPLLGSGKYGASPLCLPFGEP 720

Qy 721 STMGYVVALIILNLSLCFLMMTIAVTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780

Db 721 STMGYVVALIILNLSLCFLMMTIAVTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780

Qy 781 PVAFSLFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYLILFNPHFKEDLVSLRKQTYV 840

Db 781 PVAFSLFSSLINLTFISPEVIKFIILLVVVPLPACLNPLLYLILFNPHFKEDLVSLRKQTYV 840

Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITTYDLPPSSVPSPAYPVTESCHLSS 900

Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITTYDLPPSSVPSPAYPVTESCHLSS 900

Qy 901 VAFVPCPL 907

Db 901 VAFVPCPL 907

RESULT 4

US-10-225-567A-422

; Sequence 422, Application US/10225567A

; Publication No. US20030113798A1

; GENERAL INFORMATION:

; APPLICANT: Lifespan Biosciences

; APPLICANT: Brown, Joseph P.

; APPLICANT: Burner, Glenn C.

; APPLICANT: Roush, Christine L.

; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS

; FILE REFERENCE: 1920-4-4

; CURRENT APPLICATION NUMBER: US/10/225,567A

; CURRENT FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/257,144

; PRIOR FILING DATE: 2000-12-19

; NUMBER OF SEQ ID NOS: 2292

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 422

; LENGTH: 907

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-225-567A-422

Query Match 100.0%; Score 4702; DB 14; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSRGLVLLSLPVLQLATGSSPRSGVLLRGCPCHCEPDGGRMLLRVDCSDLGSEL 60

Db 1 MDTSRGLVLLSLPVLQLATGSSPRSGVLLRGCPCHCEPDGGRMLLRVDCSDLGSEL 60

Qy 61 PSNLVSFTSYDLDSNNNISQLLPNPLPSLRFLERLGNALTYIPKGAFTGLYSLKVL 120

Db 61 PSNLVSFTSYDLDSNNNISQLLPNPLPSLRFLERLGNALTYIPKGAFTGLYSLKVL 120

Qy 121 LQNNQLRHVPTEALQNLRLDANHI SYVPPSCFSGLSRLHRLWLDNLTETIPVQ 180

Db 121 LQNNQLRHVPTEALQNLRLDANHI SYVPPSCFSGLSRLHRLWLDNLTETIPVQ 180

Qy 181 AFRSLSALOQMTALNKTHIIPDYAFGNLSLVHLHNNRIHSLGKCFDGLHSLTLD 240

Db 181 AFRSLSALOQMTALNKTHIIPDYAFGNLSLVHLHNNRIHSLGKCFDGLHSLTLD 240

QY 241 LNYNLDLDEFTAIRTLNKLGHSSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
DB 241 LNYNLDLDEFTAIRTLNKLGHSSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
QY 301 FOHLPELRTLTLNGASQITTEFPDITGTANLESLLTGAQISSLPQTCVNCQPNLQVLDLS 360
DB 301 FOHLPELRTLTLNGASQITTEFPDITGTANLESLLTGAQISSLPQTCVNCQPNLQVLDLS 360
QY 361 YNLELDPSPSVCKQLOKIDLRHNEIYEIKVDITFQQLSLRSLNLANWKIAIHPNAPST 420
DB 361 YNLELDPSPSVCKQLOKIDLRHNEIYEIKVDITFQQLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKDLSSNLLSPFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPIAYOCC 480
DB 421 LPSLIKDLSSNLLSPFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPIAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFFEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFFEDLKALHSVQ 540
QY 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVLAGVDAFTGSPARHGAWWENGVGCHVIGFLSTIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSASVLAGVDAFTGSPARHGAWWENGVGCHVIGFLSTIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSAKFETKAPSSLKVIILLCALALATMAAVPLLGSKYKYGASPLCLPLPGE 720
DB 661 ERGFSVKYSAKFETKAPSSLKVIILLCALALATMAAVPLLGSKYKYGASPLCLPLPGE 720
QY 721 STMGYMWALILLNSLCFLMWTIAYTKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
DB 721 STMGYMWALILLNSLCFLMWTIAYTKLYCNLDKGDLENIMDCSMVKHIALLLFTNCILNC 780
QY 781 PVAFLSPSLNLTAFISPEVIFKILLVVVPLACINPLLYTLFNPHEKEDLVSRKQTYV 840
DB 781 PVAFLSPSLNLTAFISPEVIFKILLVVVPLACINPLLYTLFNPHEKEDLVSRKQTYV 840
QY 841 WTRSKHPSLMSINDSDVEKQSCDSTQALVTFTSSITVDLPSPSSVPSPAYPVTESCHLSS 900
DB 841 WTRSKHPSLMSINDSDVEKQSCDSTQALVTFTSSITVDLPSPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPCPL 907
DB 901 VAFVPCPL 907

RESULT 5

US-10-295-027-849
; Sequence 849, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666

; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 849
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-849

Query Match 100.0%; Score 4702; DB 15; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSRLGVLSSLPVLLQLATGGSPRGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
DB 1 MDTSRLGVLSSLPVLLQLATGGSPRGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
QY 61 PSNLVFTSYLDLSMNNISQLPNPLPSLRFLESLRAGNALTYIPKGAFTGLYSKVL 120
DB 61 PSNLVFTSYLDLSMNNISQLPNPLPSLRFLESLRAGNALTYIPKGAFTGLYSKVL 120
QY 121 LONNQLRHVTEALONLRSLQSLRLDANHSYVPPSCFSGHLSRLHLLWDDNALTPEVQ 180
DB 121 LONNQLRHVTEALONLRSLQSLRLDANHSYVPPSCFSGHLSRLHLLWDDNALTPEVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLTLD 240
DB 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLTLD 240
QY 241 LNYNLDLDEFTAIRTLNKLGHSSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
DB 241 LNYNLDLDEFTAIRTLNKLGHSSNNIRSIPEKAFVGNPSLTIHFYDNPFIQVGRSA 300
QY 301 FOHLPELRTLTLNGASQITTEFPDITGTANLESLLTGAQISSLPQTCVNCQPNLQVLDLS 360
DB 301 FOHLPELRTLTLNGASQITTEFPDITGTANLESLLTGAQISSLPQTCVNCQPNLQVLDLS 360
QY 361 YNLELDPSPSVCKQLOKIDLRHNEIYEIKVDITFQQLSLRSLNLANWKIAIHPNAPST 420
DB 361 YNLELDPSPSVCKQLOKIDLRHNEIYEIKVDITFQQLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKDLSSNLLSPFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPIAYOCC 480
DB 421 LPSLIKDLSSNLLSPFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPIAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFFEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFFEDLKALHSVQ 540
QY 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVLAGVDAFTGSPARHGAWWENGVGCHVIGFLSTIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSASVLAGVDAFTGSPARHGAWWENGVGCHVIGFLSTIFASESSVFLTLAAL 660

Qy	661	ERGSVKYSAKFTKAP	PSSLVKVIIL	CALLATWAAVPL	LGSKYVGASPL	CLPLPGSP	720
Db	661	ERGSVKYSAKFTKAP	PSSLVKVIIL	CALLATWAAVPL	LGSKYVGASPL	CLPLPGSP	720
Qy	721	STMGVMALII	LNSLCFLM	MTIATKLYCN	LDKDENIWD	CSMWKHIALLFT	NCILNC 780
Db	721	STMGVMALII	LNSLCFLM	MTIATKLYCN	LDKDENIWD	CSMWKHIALLFT	NCILNC 780
Qy	781	PVAFI	SPSSLN	LPISPEVI	KFILLVVVPL	PACLNPLLYIL	FNFPHFKEDVSLRKQTV 840
Db	781	PVAFI	SPSSLN	LPISPEVI	KFILLVVVPL	PACLNPLLYIL	FNFPHFKEDVSLRKQTV 840
Qy	841	WTRSKHPS	LMGINS	DDVEKQ	SCDSQALVTF	TSSSITVDL	PPSSVPSPAYPVTESCHLSS 900
Db	841	WTRSKHPS	LMGINS	DDVEKQ	SCDSQALVTF	TSSSITVDL	PPSSVPSPAYPVTESCHLSS 900
Qy	901	VAFV	PCL	907			
Db	901	VAFV	PCL	907			

RESULT 6
US-10-295-027-946
; Sequence 946, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 946
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-946

Query Match	100.0%	Score 4702;	DB 15;	Length 907;
Best Local Similarity	100.0%	Pred. No. 0;		

Matches	907;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0
Qy	1	MDTSRLGVLLSLPVLQLATGSSPSRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLS	EL	60					
Db	1	MDTSRLGVLLSLPVLQLATGSSPSRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLS	EL	60					
Qy	61	PSNLSVPTSYLDLSMNNISOLLNPLPSLAPFLBELRLAGNALTVIPKGAFTGLYS	LKVM	120					
Db	61	PSNLSVPTSYLDLSMNNISOLLNPLPSLAPFLBELRLAGNALTVIPKGAFTGLYS	LKVM	120					
Qy	121	LQNQLRHVPTALQNLRSLQSLRLDANHSIYVPPSCFSGLHSLRLHLWDDNALTEI	PVQ	180					
Db	121	LQNQLRHVPTALQNLRSLQSLRLDANHSIYVPPSCFSGLHSLRLHLWDDNALTEI	PVQ	180					
Qy	181	AFRSLSAQWMTALANKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHS	LTLD	240					
Db	181	AFRSLSAQWMTALANKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHS	LTLD	240					
Qy	241	LYNNNLDEFPFARTLSNLKELGFHSNNIIRSIPEKAFVGNPSLTIHFYDNP	QFVGRSA	300					
Db	241	LYNNNLDEFPFARTLSNLKELGFHSNNIIRSIPEKAFVGNPSLTIHFYDNP	QFVGRSA	300					
Qy	301	FQHLPELRTLTNGASQITTEFPDLTGCTANLESLTLTGAQISSLPQTCNQLP	NLQVLDLS	360					
Db	301	FQHLPELRTLTNGASQITTEFPDLTGCTANLESLTLTGAQISSLPQTCNQLP	NLQVLDLS	360					
Qy	361	YNLLEDLPSPSCVKQKIDIRNEYIYKVDTFQQLLSRLSLANWKAITHPNA	FST	420					
Db	361	YNLLEDLPSPSCVKQKIDIRNEYIYKVDTFQQLLSRLSLANWKAITHPNA	FST	420					
Qy	421	LPSLIKDLTSSNLSGFPITGLHGLTHLKTGNHALQSLISSENFPELKV	IEPVA	480					
Db	421	LPSLIKDLTSSNLSGFPITGLHGLTHLKTGNHALQSLISSENFPELKV	IEPVA	480					
Qy	481	AFVGCENAYKISQNWKNKGDNSSMDLHKDKAGMFQAQDERDLEDFLD	FEEDLKALHSVQ	540					
Db	481	AFVGCENAYKISQNWKNKGDNSSMDLHKDKAGMFQAQDERDLEDFLD	FEEDLKALHSVQ	540					
Qy	541	CSPSPGPFKECHLLDGWLLIRIGWITAVLALTCNALVTSTVRSPLIY	ISPIKLLIGVIA	600					
Db	541	CSPSPGPFKECHLLDGWLLIRIGWITAVLALTCNALVTSTVRSPLIY	ISPIKLLIGVIA	600					
Qy	601	AVNMLTGVSSAVLAGVDVAFGSGFARHGAWENGVGCHVIGFISIPASE	SSVFLTLAAL	660					
Db	601	AVNMLTGVSSAVLAGVDVAFGSGFARHGAWENGVGCHVIGFISIPASE	SSVFLTLAAL	660					
Qy	661	ERGFSVKYSAKFTKAPFSSLKVIILLCALLATMAAVPLLGSGSKYGAS	PLCLPLPFGEP	720					
Db	661	ERGFSVKYSAKFTKAPFSSLKVIILLCALLATMAAVPLLGSGSKYGAS	PLCLPLPFGEP	720					
Qy	721	STNGYVVALTILNLSLFLMWTIATKLYCNLDKGDLENIWDCSMVKHIAL	LLFTNCTILNC	780					
Db	721	STNGYVVALTILNLSLFLMWTIATKLYCNLDKGDLENIWDCSMVKHIAL	LLFTNCTILNC	780					
Qy	781	PVAPLGFSSLINITFTSPEVIKFIILVVVPLPACLNPLLYILFNPHF	KEDLVSLRKQTYV	840					
Db	781	PVAPLGFSSLINITFTSPEVIKFIILVVVPLPACLNPLLYILFNPHF	KEDLVSLRKQTYV	840					
Qy	841	WTRSKHPSLMSINSSDVEKQCDSTQALVTFTSSITYDLPPSSVPSP	APVVTESCHLSS	900					
Db	841	WTRSKHPSLMSINSSDVEKQCDSTQALVTFTSSITYDLPPSSVPSP	APVVTESCHLSS	900					
Qy	901	VAFVPC	907						
Db	901	VAFVPC	907						

RESULT 7
US-10-295-027-1331
; Sequence 1331, Application US/10295027
; Publication No. US2003023235A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel

APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
FILE OF INVENTION: Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 1331
LENGTH: 907
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-1331

Query Match 100.0%; Score 4702; DB 15; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLGVLSPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
Db 1 MDTSLGVLSPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

Qy 61 PSNLSVFTSYDLSNMNLSQLLPDLPSLRFLBELRLAGNALTYIPKGAFTGLSKVL 120
Db 61 PSNLSVFTSYDLSNMNLSQLLPDLPSLRFLBELRLAGNALTYIPKGAFTGLSKVL 120

Qy 121 LQNNQLRHVPTEALQNLRSLSLRDANHSVPPSCFSGLSRLHRLWLDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSLRDANHSVPPSCFSGLSRLHRLWLDNALTEIPVQ 180

Qy 181 AFRSLSALQAMTLALNKHHIPDYAFGNLSLVLHNNRIHSIGKCKPDLGHSLETLD 240
Db 181 AFRSLSALQAMTLALNKHHIPDYAFGNLSLVLHNNRIHSIGKCKPDLGHSLETLD 240

Qy 241 LNYNLDLFPFAIRTLNSLKLGHSHNNIRSIPEKAFVGNFSLTIHFYDNPQVGRSA 300
Db 241 LNYNLDLFPFAIRTLNSLKLGHSHNNIRSIPEKAFVGNFSLTIHFYDNPQVGRSA 300

Qy 301 FOHLPELRLTLNGASQITTEPDLTGTANLSLTLTGAISSLPTVCNQLPNLQVLDLS 360
Db 301 FOHLPELRLTLNGASQITTEPDLTGTANLSLTLTGAISSLPTVCNQLPNLQVLDLS 360

Qy 361 YNLLDLPSFVCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAFST 420

Db 361 YNLLDLPSFVCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAFST 420

Qy 421 LPSLIKLDLSSNLLSPFITGLHGLTHLKTGNHALQSLSSSENFPELKVEMPYAQQC 480
Db 421 LPSLIKLDLSSNLLSPFITGLHGLTHLKTGNHALQSLSSSENFPELKVEMPYAQQC 480

Qy 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEFLDFFEDLKALHSVQ 540

Qy 541 CSPSPGPKCEHLLDGLWLRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKCEHLLDGLWLRIGVMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Qy 601 AVNMLTGVSNAVLGAVDAFTFGSPARHGAWENGVCCHVIGFLSIFASESVFLLTAAAL 660
Db 601 AVNMLTGVSNAVLGAVDAFTFGSPARHGAWENGVCCHVIGFLSIFASESVFLLTAAAL 660

Qy 661 ERGFSVKYSKAFETKAPFSSLKVILLCALLALTAALTAALTAALTAALTAALTAAL 720
Db 661 ERGFSVKYSKAFETKAPFSSLKVILLCALLALTAALTAALTAALTAALTAALTAAL 720

Qy 721 STMGYVALIILNSLCLFLMMTIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNSLCLFLMMTIAYTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780

Qy 781 PVAFLSFSSLINLTFISPEVTKFLLVVPVLPACINPLLYLLFNPHKEDLVSRKQTYV 840
Db 781 PVAFLSFSSLINLTFISPEVTKFLLVVPVLPACINPLLYLLFNPHKEDLVSRKQTYV 840

Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSSYTDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSSYTDLPSSVPSPAYPVTESCHLSS 900

Qy 901 VAFVPC 907
Db 901 VAFVPC 907

RESULT 8

US-10-751-736-84
Sequence 84, Application US/10751736
Publication No. US20040265230A1
GENERAL INFORMATION:
APPLICANT: Wveth
APPLICANT: Martinez, Robert
APPLICANT: Brown, Eugene
APPLICANT: Liu, Wei
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
FILE OF INVENTION: CANCERS
FILE REFERENCE: AM100927 (031896-002000)
CURRENT APPLICATION NUMBER: US/10/751,736
CURRENT FILING DATE: 2003-01-06
PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
PRIOR FILING DATE: 2003-01-06
NUMBER OF SEQ ID NOS: 54873
SOFTWARE: Patent In version 3.2
SEQ ID NO 84
LENGTH: 907
TYPE: PRT
ORGANISM: Homo sapiens
US-10-751-736-84

Query Match 100.0%; Score 4702; DB 16; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLGVLSPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
Db 1 MDTSLGVLSPVLLQATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60

Qy 61 PSNLSVFTSYDLSNMNLSQLLPDLPSLRFLBELRLAGNALTYIPKGAFTGLSKVL 120

Self

Db 61 PSNLSVFTSYLDLNMNLSQLLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSIKVLM 120
QY 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSIVVPSCFSGLSLRLHMLDNNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSIVVPSCFSGLSLRLHMLDNNALTEIPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHLHNNRTHSLGKKCFDGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHLHNNRTHSLGKKCFDGLHSLETL 240
QY 241 LNNYNNLDEFPPTAIRTLNLSKELGFHNNIRSIPEKAFVGNPSLIITHFYDNPQVGRSA 300
Db 241 LNNYNNLDEFPPTAIRTLNLSKELGFHNNIRSIPEKAFVGNPSLIITHFYDNPQVGRSA 300
QY 301 FOHLPELRTLTLNGASQITEFPDLTGANLESITLTGAQISSLPQVTCNQLPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITEFPDLTGANLESITLTGAQISSLPQVTCNQLPNLQVLDLS 360
QY 361 YNLEEDLPFSVCOKLQKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHHNAPST 420
Db 361 YNLEEDLPFSVCOKLQKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHHNAPST 420
QY 421 LPSLKLDSLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPMPAYQCC 480
Db 421 LPSLKLDSLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDQERDLEDLDPFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDQERDLEDLDPFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCHEHLLDGMILIRIGWMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCHEHLLDGMILIRIGWMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCCHVIGFSLIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCCHVIGFSLIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSKAPETKAPFSSLKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGBP 720
Db 661 ERGFSVKYSKAPETKAPFSSLKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGBP 720
QY 721 STMGYMWALILNLSLCFLMWTIATYKLYCNLDKGLDENIWDSCMWKHIALLFTNCILNC 780
Db 721 STMGYMWALILNLSLCFLMWTIATYKLYCNLDKGLDENIWDSCMWKHIALLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFTISPEVIRKIFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTV 840
Db 781 PVAFLSFSSLINLTFTISPEVIRKIFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTV 840
QY 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFTSSSITYDLPSSVSPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFTSSSITYDLPSSVSPSPAYPVTESCHLSS 900
QY 901 VAFVPECL 907
Db 901 VAFVPECL 907

RESULT 9

US-10-482-029-158
; Sequence 158, Application US/10482029
; Publication No. US20050037445A1
; GENERAL INFORMATION:
; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482, 029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 158
; LENGTH: 907

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-482-029-158
Query Match 100.0%; Score 4702; DB 17; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLRGVLLSLPVLQALATGSSPRSGVLLRGCPHCHCEPDGRMLLVDCSDGLSEL 60
Db 1 MDTSLRGVLLSLPVLQALATGSSPRSGVLLRGCPHCHCEPDGRMLLVDCSDGLSEL 60
QY 61 PSNLSVFTSYLDLNMNLSQLLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSIKVLM 120
Db 61 PSNLSVFTSYLDLNMNLSQLLPNPLPSLRFLEELRLAGNALTYIPKGAFTGLYSIKVLM 120
QY 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSIVVPSCFSGLSLRLHMLDNNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSIVVPSCFSGLSLRLHMLDNNALTEIPVQ 180
QY 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHLHNNRTHSLGKKCFDGLHSLETL 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHLHNNRTHSLGKKCFDGLHSLETL 240
QY 241 LNNYNNLDEFPPTAIRTLNLSKELGFHNNIRSIPEKAFVGNPSLIITHFYDNPQVGRSA 300
Db 241 LNNYNNLDEFPPTAIRTLNLSKELGFHNNIRSIPEKAFVGNPSLIITHFYDNPQVGRSA 300
QY 301 FOHLPELRTLTLNGASQITEFPDLTGANLESITLTGAQISSLPQVTCNQLPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITEFPDLTGANLESITLTGAQISSLPQVTCNQLPNLQVLDLS 360
QY 361 YNLEEDLPFSVCOKLQKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHHNAPST 420
Db 361 YNLEEDLPFSVCOKLQKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHHNAPST 420
QY 421 LPSLKLDSLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPMPAYQCC 480
Db 421 LPSLKLDSLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDQERDLEDLDPFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDQERDLEDLDPFEEDLKALHSVQ 540
QY 541 CSPSPGPKPCHEHLLDGMILIRIGWMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCHEHLLDGMILIRIGWMTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCCHVIGFSLIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVLAGVDAFTFGSFARHGAWENGVCCHVIGFSLIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSKAPETKAPFSSLKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGBP 720
Db 661 ERGFSVKYSKAPETKAPFSSLKVIILLCALLATMAAVPLLGSKYKYGASPLCLPLPFGBP 720
QY 721 STMGYMWALILNLSLCFLMWTIATYKLYCNLDKGLDENIWDSCMWKHIALLFTNCILNC 780
Db 721 STMGYMWALILNLSLCFLMWTIATYKLYCNLDKGLDENIWDSCMWKHIALLFTNCILNC 780
QY 781 PVAFLSFSSLINLTFTISPEVIRKIFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTV 840
Db 781 PVAFLSFSSLINLTFTISPEVIRKIFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTV 840
QY 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFTSSSITYDLPSSVSPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFTSSSITYDLPSSVSPSPAYPVTESCHLSS 900
QY 901 VAFVPECL 907
Db 901 VAFVPECL 907

RESULT 10
US-10-505-486-103
; Sequence 103, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 103
; LENGTH: 1145
; TYPE: PRT
; ORGANISM: Human
US-10-505-486-103

Query Match 100.0%; Score 4701; DB 17; Length 1145;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 906; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLQLATGGSPRSGLVLRGCPTHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSLRGVLLSLPVLQLATGGSPRSGLVLRGCPTHCHCEPDGRMLLRVDCSDLGSEL 60

Qy 61 PSNLSVFTSYLDLSMNNISQLLPNPLSLRFLLELRAGNALTYIPKGAFTGLYSLKVL 120
Db 61 PSNLSVFTSYLDLSMNNISQLLPNPLSLRFLLELRAGNALTYIPKGAFTGLYSLKVL 120

Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSYVPPSCFSGLSLRLHLDNNALETIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSYVPPSCFSGLSLRLHLDNNALETIPVQ 180

Qy 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRHSGLGKCPDGLHSLTLD 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRHSGLGKCPDGLHSLTLD 240

Qy 241 LNYNNLDEFFTAIRTLNKLGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPIDFVGSA 300
Db 241 LNYNNLDEFFTAIRTLNKLGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPIDFVGSA 300

Qy 301 FOHLPELRTLTLNGASQITTEPDLTGTLNLESRLTGTGAQISSLPQTCVNCNLPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITTEPDLTGTLNLESRLTGTGAQISSLPQTCVNCNLPNLQVLDLS 360

Qy 361 LPSLTKLDSLNNLSFFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPYVQCC 480
Db 361 LPSLTKLDSLNNLSFFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPYVQCC 480

Qy 481 AFGVCENAYKISQNNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFEEDLKALHSVQ 540

Qy 541 CSPSPGPKPEHLLDGWLIRIGVWTVIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPEHLLDGWLIRIGVWTVIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600

Qy 601 AVNMLTGVSVAVLACVDAFTGSPARHGAWENGVCVIGFLSTFASSESVLLTAAAL 660
Db 601 AVNMLTGVSVAVLACVDAFTGSPARHGAWENGVCVIGFLSTFASSESVLLTAAAL 660

Qy 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720

Db 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720
Qy 721 STMGYVALIILNSLCFLMMTIAVTKLYCNLDKGDLENINWDCSMVKHIALALLFTNCILNC 780
Db 721 STMGYVALIILNSLCFLMMTIAVTKLYCNLDKGDLENINWDCSMVKHIALALLFTNCILNC 780
Qy 781 PVAFLSFSSILNLTFFISPEVIKFILLVVVPLPACINPLLYLTFNPHFKEDIVSLRKQTYV 840
Db 781 PVAFLSFSSILNLTFFISPEVIKFILLVVVPLPACINPLLYLTFNPHFKEDIVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVTTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVTTESCHLSS 900
Qy 901 VAFVPCL 907
Db 901 VAFVPCL 907

RESULT 11
US-10-251-385-278
; Sequence 278, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 278
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-278

Query Match 99.9%; Score 4696; DB 14; Length 907;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 906; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLQLATGGSPRSGLVLRGCPTHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSLRGVLLSLPVLQLATGGSPRSGLVLRGCPTHCHCEPDGRMLLRVDCSDLGSEL 60

Qy 61 PSNLSVFTSYLDLSMNNISQLLPNPLSLRFLLELRAGNALTYIPKGAFTGLYSLKVL 120
Db 61 PSNLSVFTSYLDLSMNNISQLLPNPLSLRFLLELRAGNALTYIPKGAFTGLYSLKVL 120

Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSYVPPSCFSGLSLRLHLDNNALETIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSYVPPSCFSGLSLRLHLDNNALETIPVQ 180

Qy 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRHSGLGKCPDGLHSLTLD 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRHSGLGKCPDGLHSLTLD 240

Qy 241 LNYNNLDEFFTAIRTLNKLGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPIDFVGSA 300
Db 241 LNYNNLDEFFTAIRTLNKLGFHSHNNIRSIPEKAFVGNPSLTIHFYDNPIDFVGSA 300

Qy 301 FOHLPELRTLTLNGASQITTEPDLTGTLNLESRLTGTGAQISSLPQTCVNCNLPNLQVLDLS 360
Db 301 FOHLPELRTLTLNGASQITTEPDLTGTLNLESRLTGTGAQISSLPQTCVNCNLPNLQVLDLS 360

Qy 361 LPSLTKLDSLNNLSFFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPYVQCC 480
Db 361 LPSLTKLDSLNNLSFFITGLHGLTHLKTGNHALQSLISSENPPELKVTEMPYVQCC 480

Qy 481 AFGVCENAYKISQNNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNNKGDNSMDDLHKDKAGMFOQDERDLEDFLLDFEEDLKALHSVQ 540

Qy 541 CSPSPGPKPEHLLDGWLIRIGVWTVIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPEHLLDGWLIRIGVWTVIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600

Qy 601 AVNMLTGVSVAVLACVDAFTGSPARHGAWENGVCVIGFLSTFASSESVLLTAAAL 660
Db 601 AVNMLTGVSVAVLACVDAFTGSPARHGAWENGVCVIGFLSTFASSESVLLTAAAL 660

Qy 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP 720

Db 361 YNLEEDLPSFVCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEMPAYQCC 480
Db 421 LPSLIKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ 540
QY 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWNGVGVIGFSLIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWNGVGVIGFSLIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLGSGKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLGSGKYGASPLCLPFGEP 720
QY 721 STMGYVVALIILNSLCFLMWTIATYKLYCNLDKGLDENIWDSCMKHIALLLFTNCILNC 780
Db 721 STMGYVVALIILNSLCFLMWTIATYKLYCNLDKGLDENIWDSCMKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTIFISPEVIKFIILLVVPPLACLNPLLYIILFNPHEKEDLVSLRKQTVY 840
Db 781 PVAFLSFSSLINLTIFISPEVIKFIILLVVPPLACLNPLLYIILFNPHEKEDLVSLRKQTVY 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPECL 907
Db 901 VAFVPECL 907
RESULT 12
US-09-970-944-27
; Sequence 27, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shinkens, Richard A
; TITLE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; PRIOR FILING DATE: 2002-05-02
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 27
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-27
Query Match 99.8%; Score 4691; DB 10; Length 907;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 905; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGMLLRVDCSDGLSEL 60
Db 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGMLLRVDCSDGLSEL 60
QY 61 PSLNSVFTSYDLSNMNLSQLLPNPLPSLFLFEEELRAGNALTYIPKGAFTGLYSKVLV 120
Db 61 PSLNSVFTSYDLSNMNLSQLLPNPLPSLFLFEEELRAGNALTYIPKGAFTGLYSKVLV 120

QY 121 LQNNQLRHVPTALQNLRSLOSLRDANHSYVPPSCFSGLSLRLHMLDNDALTEIPVQ 180
Db 121 LQNNQLRHVPTALQNLRSLOSLRDANHSYVPPSCFSGLSLRLHMLDNDALTEIPVQ 180
QY 181 AFRSLSALQAMTLALANKIHHIPDYAFGNLSLVLHLHNNRTHSLGKCKFDGLHSLTLD 240
Db 181 AFRSLSALQAMTLALANKIHHIPDYAFGNLSLVLHLHNNRTHSLGKCKFDGLHSLTLD 240
QY 241 LNNYNNLDFPPTAIRTLNKLPGFHSNNIRSIPEKAFVGNPNSLIITTHFYDNPFIQFVGSA 300
Db 241 LNNYNNLDFPPTAIRTLNKLPGFHSNNIRSIPEKAFVGNPNSLIITTHFYDNPFIQFVGSA 300
QY 301 FOHLPELRTLTLNGASQITEFPDLTGTANLESILTGTAGQISSLPQTVCNQNLNQLVLDLS 360
Db 301 FOHLPELRTLTLNGASQITEFPDLTGTANLESILTGTAGQISSLPQTVCNQNLNQLVLDLS 360
QY 361 YNLEEDLPSFVCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAPST 420
Db 361 YNLEEDLPSFVCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEMPAYQCC 480
Db 421 LPSLIKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEMPAYQCC 480
QY 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ 540
QY 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWNGVGVIGFSLIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSAAVLAVDAFTFGSFARHGAWNGVGVIGFSLIFASESSVFLTLAAL 660
QY 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLGSGKYGASPLCLPFGEP 720
Db 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLGSGKYGASPLCLPFGEP 720
QY 721 STMGYVVALIILNSLCFLMWTIATYKLYCNLDKGLDENIWDSCMKHIALLLFTNCILNC 780
Db 721 STMGYVVALIILNSLCFLMWTIATYKLYCNLDKGLDENIWDSCMKHIALLLFTNCILNC 780
QY 781 PVAFLSFSSLINLTIFISPEVIKFIILLVVPPLACLNPLLYIILFNPHEKEDLVSLRKQTVY 840
Db 781 PVAFLSFSSLINLTIFISPEVIKFIILLVVPPLACLNPLLYIILFNPHEKEDLVSLRKQTVY 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPECL 907
Db 901 VAFVPECL 907

RESULT 13
US-10-295-027-484
; Sequence 484, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.

;; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
;; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
;; FILE REFERENCE: 018501-012500US
;; CURRENT APPLICATION NUMBER: US/10/295,027
;; CURRENT FILING DATE: 2002-11-13
;; PRIOR APPLICATION NUMBER: US 09/663,733
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: US 60/350,666
;; PRIOR FILING DATE: 2001-11-13
;; PRIOR APPLICATION NUMBER: US 60/335,394
;; PRIOR FILING DATE: 2001-11-15
;; PRIOR APPLICATION NUMBER: US 60/332,464
;; PRIOR FILING DATE: 2001-11-21
;; PRIOR APPLICATION NUMBER: US 60/334,393
;; PRIOR FILING DATE: 2001-11-29
;; PRIOR APPLICATION NUMBER: US 60/340,376
;; PRIOR FILING DATE: 2001-12-14
;; PRIOR APPLICATION NUMBER: US 60/347,211
;; PRIOR FILING DATE: 2002-01-08
;; PRIOR APPLICATION NUMBER: US 60/347,349
;; PRIOR FILING DATE: 2002-01-10
;; PRIOR APPLICATION NUMBER: US 60/355,250
;; PRIOR FILING DATE: 2002-02-08
;; PRIOR APPLICATION NUMBER: US 60/356,714
;; PRIOR FILING DATE: 2002-02-13
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 1386
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 484
;; LENGTH: 883
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-295-027-484

Query Match 96.9%; Score 4556; DB 15; Length 883;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 883; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

Qy 1 MDTSLGVLISLPLVLLQATGSSPRSGVILRGCTHCHCEPDGMLRVDCSDLGSEL 60
Db 1 MDTSLGVLISLPLVLLQATGSSPRSGVILRGCTHCHCEPDGMLRVDCSDLGSEL 60
Qy 61 PSNLSVFTSYLDLNNISQLPNPLSLRFLLELRAGNALTYPKGAFTGLSLKVL 120
Db 61 PSNLSVFTSYLDLNNISQLPNPLSLRFLLELRAGNALTYPKGAFTGLSLKVL 120
Qy 121 LQNNQLRHVPTEALQNLRSLSRLDANHSYVPPSCFSLHSLRHLWLDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSRLDANHSYVPPSCFSLHSLRHLWLDNALTEIPVQ 180
Qy 181 AFRSLALQMTLALNKIHHIPDYAFGNLSLVLHLHNNRIHSLGKCFGLHSLETL 240
Db 181 AFRSLALQMTLALNKIHHIPDYAFGNLSLVLHLHNNRIHSLGKCFGLHSLETL 240
Qy 241 LNNYNDLDEFTAIRTLNKLGLPHSNNIRISPEKAFVGNPFLTIHFYDNPQVGRSA 300
Db 241 LNNYNDLDEFTAIRTLNKLGLPHSNNIRISPEKAFVGNPFLTIHFYDNPQVGRSA 276
Qy 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLPTVCNQLPNLQVLDLS 360
Db 277 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLPTVCNQLPNLQVLDLS 336
Qy 361 YNLLDLPSFVSCQKLOKIDIRHNEIYEIKVDYTFQOILSLRSLNLANWKIHHNPEST 420
Db 337 YNLLDLPSFVSCQKLOKIDIRHNEIYEIKVDYTFQOILSLRSLNLANWKIHHNPEST 396
Qy 421 LPSLTKLDSNLLSSFFITGLHGLTHLKTGNHALQSLISSENPELKVIEPMPYQCC 480
Db 397 LPSLTKLDSNLLSSFFITGLHGLTHLKTGNHALQSLISSENPELKVIEPMPYQCC 456
Qy 481 AFGVCENAYKISNQWNGDSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 457 AFGVCENAYKISNQWNGDSSMDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 516

Qy 541 CSPSPGPKPCHEHLIDGWLIRIGVTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 517 CSPSPGPKPCHEHLIDGWLIRIGVTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 576
Qy 601 AVNMLTGVSASVLAGVDAFTFGSPARHGAWWGVCHVIGFLSIFASESVFLLTLAAL 660
Db 577 AVNMLTGVSASVLAGVDAFTFGSPARHGAWWGVCHVIGFLSIFASESVFLLTLAAL 636
Qy 661 ERGSVKYSKAFETKAPFSSKVLILLCALLATMAAVPLLGSKYKASPLCLPLPGE 720
Db 637 ERGSVKYSKAFETKAPFSSKVLILLCALLATMAAVPLLGSKYKASPLCLPLPGE 696
Qy 721 STMGVMVALILLNSLCFLMMTIAVTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 697 STMGVMVALILLNSLCFLMMTIAVTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 756
Qy 781 PVAFLSFSSILNLTFTISPEVIKFI LLVVVPLPACLNPLLYILFNPHFKEDLVSRKQTV 840
Db 757 PVAFLSFSSILNLTFTISPEVIKFI LLVVVPLPACLNPLLYILFNPHFKEDLVSRKQTV 816
Qy 841 WTRSKHPSLMSINSDVDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYVPTESCHLSS 900
Db 817 WTRSKHPSLMSINSDVDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYVPTESCHLSS 876
Qy 901 VAFVPC 907
Db 877 VAFVPC 883

RESULT 14
US-10-295-027-1330
; Sequence 1330, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1


```

; SEQ ID NO 1330
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1330

Query Match      96.9%; Score 4556; DB 15; Length 883;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 883; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 MDTSLGVLSPVLVQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB 1 MDTSLGVLSPVLVQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
QY 61 PNLNVFTSYLDLNMNNSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
DB 61 PNLNVFTSYLDLNMNNSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
QY 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSLRLHLLWDDNALTEIPVQ 180
DB 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSLRLHLLWDDNALTEIPVQ 180
QY 181 AFRSLSALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRTHSLGKCFDGLHSL 240
DB 181 AFRSLSALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRTHSLGKCFDGLHSL 240
QY 241 LNNNLDPEPTAIRTLNKLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNP 300
DB 241 LNNNLDPEPTAIRTLNKLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNP 300
QY 301 FOHLPELRTLTLNGASQITEFPDLTGTLNLESILTLTGAQISSLPQTCVQNL 360
DB 277 FOHLPELRTLTLNGASQITEFPDLTGTLNLESILTLTGAQISSLPQTCVQNL 360
QY 361 YNLLEDLPSPVSCQKQKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKI 420
DB 337 YNLLEDLPSPVSCQKQKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKI 420
QY 421 LPSLKLKDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKV 480
DB 397 LPSLKLKDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKV 480
QY 481 AFGVCENAYKISNQWKNKGNSSMDLHKDAGMFOQDERDLEDFLLDP 540
DB 457 AFGVCENAYKISNQWKNKGNSSMDLHKDAGMFOQDERDLEDFLLDP 540
QY 541 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPL 600
DB 517 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPL 600
QY 601 AVNMLTGVSAAVLGADFTFGSFARHGAWENGCHVIGFLSIFASESSV 660
DB 577 AVNMLTGVSAAVLGADFTFGSFARHGAWENGCHVIGFLSIFASESSV 660
QY 661 ERGFSVKYSKAPETKAPSSSLKVIILLCALLATWAAVPLLGGSKYGAS 720
DB 637 ERGFSVKYSKAPETKAPSSSLKVIILLCALLATWAAVPLLGGSKYGAS 720
QY 721 STMGYVAIILANSICFLMWTIATKLYCNLDKGLNWDSCMWKHTALL 780
DB 697 STMGYVAIILANSICFLMWTIATKLYCNLDKGLNWDSCMWKHTALL 780
QY 781 PVAFLSFSSLINLTFISPEVIKILLVVPVLPACLNPLLYILFNP 840
DB 757 PVAFLSFSSLINLTFISPEVIKILLVVPVLPACLNPLLYILFNP 840
QY 841 WTRSKHPSLMSINSDVEKQSCDSTQALVTFSSSTIYDLPSSVPSP 900
DB 817 WTRSKHPSLMSINSDVEKQSCDSTQALVTFSSSTIYDLPSSVPSP 900
QY 901 VAFVPC 907
DB 877 VAFVPC 883
```

RESULT 15

```

US-10-173-999-28
; Sequence 28, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 28
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-28

Query Match      96.9%; Score 4556; DB 15; Length 883;
Best Local Similarity 97.4%; Pred. No. 0;
Matches 883; Conservative 0; Mismatches 0; Indels 24; Gaps 1;

QY 1 MDTSLGVLSPVLVQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB 1 MDTSLGVLSPVLVQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
QY 61 PNLNVFTSYLDLNMNNSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
DB 61 PNLNVFTSYLDLNMNNSQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVL 120
QY 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSLRLHLLWDDNALTEIPVQ 180
DB 121 LQNNQLRHVPTALQNLRSLSRLDANHSYVPPSCFSGLSLRLHLLWDDNALTEIPVQ 180
QY 181 AFRSLSALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRTHSLGKCFDGLHSL 240
DB 181 AFRSLSALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRTHSLGKCFDGLHSL 240
QY 241 LNNNLDPEPTAIRTLNKLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNP 300
DB 241 LNNNLDPEPTAIRTLNKLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNP 300
QY 301 FOHLPELRTLTLNGASQITEFPDLTGTLNLESILTLTGAQISSLPQTCVQNL 360
DB 277 FOHLPELRTLTLNGASQITEFPDLTGTLNLESILTLTGAQISSLPQTCVQNL 360
QY 361 YNLLEDLPSPVSCQKQKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKI 420
DB 337 YNLLEDLPSPVSCQKQKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKI 420
QY 421 LPSLKLKDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKV 480
DB 397 LPSLKLKDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELKV 480
QY 481 AFGVCENAYKISNQWKNKGNSSMDLHKDAGMFOQDERDLEDFLLDP 540
DB 457 AFGVCENAYKISNQWKNKGNSSMDLHKDAGMFOQDERDLEDFLLDP 540
QY 541 CSPSPGPPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPL 600
```


Db	517		CSPSPGPKCEHLLDGLIIRIGVWTIAVLATCNALVTSVFRSPLYISPIKLLIGVIA	576
Qy	601		AVNMLTGVSSAVLAGVDAFTFGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL	660
Db	577		AVNMLTGVSSAVLAGVDAFTFGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL	636
Qy	661		ERGFVKYSKAFETKAPFSSIKVILLCALLALTMAAVPLLGSSKYGASPLCLPLPGEP	720
Db	637		ERGFVKYSKAFETKAPFSSIKVILLCALLALTMAAVPLLGSSKYGASPLCLPLPGEP	696
Qy	721		STMGYVALIILNSLCFLMWTIAVTKLYCNLDKGDLENINWDCSMVKHIALLLFTNCILNC	780
Db	697		STMGYVALIILNSLCFLMWTIAVTKLYCNLDKGDLENINWDCSMVKHIALLLFTNCILNC	756
Qy	781		PVAFLSFSSLINLTFISPEVIKFILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYV	840
Db	757		PVAFLSFSSLINLTFISPEVIKFILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYV	816
Qy	841		WTRSKHPSLMSINSDDEKQSCDSTQALVTFSSITTYDLPPSSVPSPAYPVTESCHLSS	900
Db	817		WTRSKHPSLMSINSDDEKQSCDSTQALVTFSSITTYDLPPSSVPSPAYPVTESCHLSS	876
Qy	901		VAFVPCL 907	
Db	877		VAFVPCL 883	

Search completed: July 12, 2005, 07:48:45
Job time : 621 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2005, 07:51:27 ; Search time 1145 Seconds
(without alignments)
4689.262 Million cell updates/sec

Title: US-10-751-736-84

Perfect score: 907

Sequence: 1 MDTSRLGVLSPVLLQLAT.....PAYPTESCHLSSVAFVPCFL 907

Scoring table:

OLIGO 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Word size: 1

Total number of hits satisfying chosen parameters: 8776198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cpn2.1/USPTO.spool/US10751736/runat.12072005.083640.22162/app.query.fasta.1.1095
-DB=N_Geneseq.15Dec04 -QFMT=fastap -SUPFIX=oli.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751736 @CGN 1.1.810 @runat.12072005.083640.22162 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPHLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N_Geneseq.16Dec04:*

1: Geneseq1980s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	2724	3 AAA30770	AAA30770 Human G p
2	907	100.0	2724	8 ABZ42816	ABZ42816 Human G p
3	907	100.0	2724	10 ADC22782	ADC22782 Human G p
4	907	100.0	2724	10 ADH14255	ADH14255 Human HG3
5	907	100.0	2724	12 ADI32985	ADI32985 Human G p

6	907	100.0	2724	12	ADO29923	Ado29923 Human GPC
7	907	100.0	2880	11	ADN39796	Adn39796 Cancer/an
8	907	100.0	2880	13	ADQ80249	Adq80249 G protein
9	907	100.0	2973	13	ADR67869	Adr67869 Human HG3
10	907	100.0	3032	11	ADN39627	Adn39627 Cancer/an
11	907	100.0	3032	11	ADN39530	Adn39530 Cancer/an
12	907	100.0	4558	2	AAZ23980	Aaz23980 Human HG3
13	907	100.0	4570	12	ADL12472	Adl12472 Human ste
14	833	91.8	3297	2	AAA323981	Aaa323981 Human HG3
15	820	90.4	2724	3	AAA30779	Aaa30779 DNA encod
16	820	90.4	2724	10	ADC22796	Adc22796 Human G p
17	820	90.4	2724	10	ADH14269	Adh14269 Mutated h
18	806	88.9	3438	10	ADF70582	Adf70582 Orphan re
19	621	68.5	2651	10	ADB80463	Adb80463 Ovarian c
20	621	68.5	2651	11	ADN39165	Adn39165 Cancer/an
21	621	68.5	2651	11	ADN39795	Adn39795 Cancer/an
22	606	66.8	2082	2	AAZ25344	Aaz25344 Human LGR
23	232	25.6	723	8	ACA04772	ACA04772 cDNA enco
24	202	22.3	606	2	AZA40459	Aza40459 Human gon
25	197	21.7	1790	6	ABQ55077	Abq55077 Human ova
26	51	5.6	2724	12	ADR30213	Adr30213 Mouse GPC
27	51	5.6	3098	13	ADR67870	Adr67870 Mouse HG3
28	23	2.5	2636	5	ABA09697	Abao9697 Human bon
29	19	2.1	60	6	ABN47039	Abn47039 Human spl
30	18	2.0	284	2	AAZ40463	Aaz40463 Human gon
31	18	2.0	798	3	AAC75792	Aac75792 Human ORF
32	18	2.0	1827	12	ADQ64302	Adq64302 Novel hum
33	18	2.0	2208	6	AAI67925	Aai67925 Human LGR
34	18	2.0	2208	11	ADN02247	Adn02247 Human par
35	18	2.0	2208	12	ADK19412	Adk19412 ORF of pa
36	18	2.0	2376	12	ADO29126	Ado29126 Mouse nov
37	18	2.0	2484	12	ADO29786	Ado29786 Human nov
38	18	2.0	2484	12	ADQ76788	Adq76788 Leucine-r
39	18	2.0	2487	6	ADH50804	Adh50804 Human G-p
40	18	2.0	2711	6	AAI67924	Aai67924 Partial c
41	18	2.0	2711	11	ADN02245	Adn02245 Human par
42	18	2.0	2711	12	ADK19410	Adk19410 Partial c
43	18	2.0	2786	10	ADC16710	Adc16710 Human G-p
44	18	2.0	2901	6	AAI67927	Aai67927 Human LGR
45	18	2.0	2901	6	AAI67921	Aai67921 Mouse LGR

ALIGNMENTS

RESULT 1
AAA30770
ID AAA30770 standard; cDNA; 2724 BP.
XX
AC AAA30770;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor HG38 cDNA.
XX
KW G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening; agonist;
KW antagonist; ss.
XX
OS Homo sapiens.
XX
PN WO200022129-A1.
XX
PD 20-APR-2000.
XX
PF 12-OCT-1999; 99WO-US023938.
XX
PR 13-OCT-1998; 98US-00170496.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Behan DP, Chalmers DT, Liaw CW;
XX
DR WPI; 2000-329165/28.

DR P-PSDB; AAY90682.

XX Non-endogenous constitutively activated human G protein-coupled

PT receptors, useful for identifying agonists for use as pharmaceutical

PT agents.

XX

PS Example 1; Page 315-317; 341pp; English.

XX

CC The invention relates to constitutively active, non-endogenous versions

CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-

CC AAY90677 and AAY90683-Y90687), and to DNA encoding them (AAA30709-A30743

CC and AAA30775-A30779). The mutant proteins of the invention contain a

CC mutation in a portion of the protein comprising intracellular loop 3

CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,

CC is substituted for an endogenous residue in IC3 at a position 16 amino

CC acids N-terminal of an endogenous proline in TM6 to form a sequence X-

CC (AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg or

CC Ala, and is preferably Lys. When the endogenous residue at this position

CC is Lys, this residue is replaced by His, Arg or preferably Ala. The 15

CC amino acid stretch between the substituted amino acid and the Pro may be

CC endogenous, non-endogenous, or a mixture of endogenous and non-endogenous

CC residues. The constitutively active GPCRs are useful for identifying

CC antagonists, agonists and partial agonists for use as pharmaceutical

CC agents. The mutant proteins are also useful in research settings for

CC elucidating the roles of the receptors in normal and diseased conditions.

CC Antagonists for a particular GPCR are useful for treating diseases and

CC disorders associated with that receptor. Because the novel mutant GPCRs

CC are constitutively active, they can be used directly for screening of

CC compounds without the need for endogenous ligands. The present sequence

CC represents cDNA encoding a human wild-type GPCR used in an

CC exemplification of the invention. This was cloned and subjected to site-

CC directed mutagenesis (SDM) to generate DNA encoding the corresponding

CC mutant of the invention

XX

SQ Sequence 2724 BP; 665 A; 723 C; 554 G; 782 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2724
Score:	907.00	Matches:	907
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-10-751-736-84 (1-907) x AAA30770 (1-2724)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuGlnLeuAlaThr 20

DB 1 ATGGACACTCCCGCTCGGTGTGCTCTGTCTTCTGCTGCTGCTGCGGACC 60

QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40

DB 61 GGGGGCAGCTCTCCAGGCTCTGGTGTGCTGAGGGGGTGGCCACACACTGTCTATG 120

QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60

DB 121 GAGCCCGCAGCGCAGATGTGCTAGGGGGGAGCTGCTCGACCTGGGGCTCTCGGAGCTG 180

QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80

DB 181 CTTTCCACCTCAGCGCTCTTCACTTCTTACCTAGACCTCAGTATGACCAACATCAGTCAG 240

QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuLeuArgLeuAlaGlyAsn 100

DB 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCGAGGAGTACGCTTCGCGGAAAC 300

QY 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120

DB 301 GCTCTGACATACATTTCCCAAGGGAGCATTCCTCTGCGCTTTACAGTCTTAAAGTCTTATG 360

QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140

DB 361 CTGCAGATATATCAGCTAAGACACAGTACCACAGAAGCTCTGCAGAAATTTGGAGGCCTT 420

QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160

DB 421 CAATCCCTGCGCTCGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480

QY 161 LeuHisSerLeuArgHisLeuTyrLeuAspAsnAlaLeuThrGluLeuProValGln 180

DB 481 CTGCATTCCTTGAGGCACCTCTGGCTGGATGCAATGCTTAAACAGAAATCCCGTCCAG 540

QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200

DB 541 GCTTTAGAGTTTATCGGCATTTGCAAGCCATGACTTGGCCCTGGAACAAATAACACCAC 600

QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220

DB 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAAACAT 660

QY 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240

DB 661 AGAATCCACTCCCTGGGAAGAATGCTTTGATGGCTCCACAGCTAGAGACTTTAGAT 720

QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260

DB 721 TTAATTAACAATAACCTTGATGAATTCGCCACTGCAATTAGGACACTCTCCAACCTTAA 780

QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280

DB 781 GAATAGGATTTCAATAGCAACAATATCAGGTGCGATCTGAGAAAGCAATTTGTAGGCAAC 840

QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300

DB 841 CTTTCTTATTACAAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 900

QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320

DB 901 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTGAATGGTGGCTCACAATACTGAA 960

QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuGlyValaGlnIle 340

DB 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020

QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360

DB 1021 TCATCTCTCTCTCAACCTGCTGCAATCAGTTACCTATCTCCAAGTCTAGATCTGCT 1080

QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380

DB 1081 TACAACCTTATTAGAAGATTTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC 1140

QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400

DB 1141 CTAGACATAATGAATCTAGAAATTAAGTTAGACATTTCCAGCAGATGTTGTTAGCCTC 1200

QY 401 ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleHisProAsnAlaPheSerThr 420

DB 1201 CGATCGCTGAATTTGGCTTGGACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260

QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440

DB 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCCAACCTCTGCTGCTCTTCTCTTAACT 1320

QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460

DB 1321 GGGTTACATGGTTTAACTCAGTTTAAAGTAAAGAAATCATGCTCTTACAGAGCTTGATA 1380

QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480

DB 1381 TCATCTGAAACTTTCAGAACTCAAGGTATAGAAATGCCTTATGCTTACCAGTCTGT 1440

QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnThrAsnLysGlyAspAsn 500

DB 1441 GCATTTGGAGTGTGTGAGAATGCCTTATAAGATTTCTTAATCAATGGAATAAAGGTGACAC 1500

QY 501 SerSerMetAspAspLeuHisIlyIysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGAGGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGACCT 1560
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLeuHisSerValGln 540
DB 1561 GACCTTGAGATTTCTGCTTGACCTTTGAGGAGACCTGAAAGCCCTTCATTCAGTGCG 1620
QY 541 CysSerProSerProGlyProPheIysProCysGluHisLeuLeuAspGlyTrpLeuile 560
DB 1621 TGTTCACCTTCCAGGCCCCCTTCAACCCCTGTAACACCTGCTTGATGGCTGGCTGATC 1680
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATTTGGAGTGTGACCATAGACATGCTGAGCACTTACTTGTAAATGCTTTGGTGAATCA 1740
QY 581 ThrValPheArgSerProLeuLeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 600
DB 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCATTAACCTGTTAATTTGGGGTCAATCGCA 1800
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGTGGATCGTTCACT 1860
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValle 640
DB 1861 TTTGGCAGCTTTTGCACACATGCTGCTGGTGGAGATGGGGTGGTGGCCATGTCATT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
DB 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTCTCTTACTCTGCGAGCCCTG 1980
QY 661 GluArgGlyPheSerValIysTyrSerAlaIysPheGluThrIysAlaPheProPheSerSer 680
DB 1981 GAGCGTGGGTCTCTGTGAATATCTGCAAAATTTGAACGAGGCTCATTTTCTAGC 2040
QY 681 LeuIysValIleIleLeuLeuCysAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 700
DB 2041 CTGAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 701 LeuGlyGlySerIysTyrGlyAlaSerProLeuLeuLeuLeuLeuLeuLeuLeuLeu 720
DB 2101 CTGGGTGGCAGCAGTATGCGCCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
DB 2161 AGCACCATGGGTACATGCTGCTCACTCTGCTCACTCTGCTCACTCTGCTCTCTCATG 2220
QY 741 ThrIleAlaTyrThrIysLeuTyrCysAsnLeuLeuLeuLeuLeuLeuLeuLeuLeu 760
DB 2221 ACCATTTGCTTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAATATTTGG 2280
QY 761 AspCysSerMetValIysHisIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 780
DB 2281 GACTGCTCTATGTTAAACACATTTGCTGCTTCTTCCCACTGCACTCTTAACTGC 2340
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
DB 2341 CTGTGGCTCTTCTGTGCTCTCTCTCTTAAATAACCTTACATTTATCAGTCTGAATGA 2400
QY 801 IleIysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
DB 2401 ATTAAGTTTATCTTCTGCTGGTAGTCCCACTTCTGCTGATGCTCAATCCCTTCTTAC 2460
QY 821 IleLeuPheAsnProHisPheIysGluAspLeuValSerLeuArgIysGlnThrIleVal 840
DB 2461 ATCTTCTTCAATCTCACTTTAAGAGGATCTGGTGGAGCTCGAAGCAAACTACGTC 2520
QY 841 TrpThrArgSerIlyshiProSerLeuMetSerIleAsnSerAspAspValGluIysGln 860
DB 2521 TGGACAAGATCAAAACCAACCAAGCTTGTATGCTCAATTAATCTGTATGATGCGAAACAG 2580
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrIleThrAspLeu 880

DB 2581 TCTGTGACTCAACTCAAGCCTTGTAACTTTTACAGCTCCAGATCACTTATGACCTG 2640
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
DB 2641 CTTCCAGTTCGTCGCCATCCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
QY 901 ValAlaPheValProCysLeu 907
DB 2701 GTGGCATTTGTCCCATGTCTC 2721

RESULT 2

ABZ42816

ID ABZ42816 standard; DNA; 2724 BP.

XX AC ABZ42816;

XX DT 04-MAR-2003 (first entry)

XX DE Human G protein-coupled receptor GPR49 nucleotide SEQ ID NO:421.

XX KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer; gene; ds.

XX OS Homo sapiens.

XX PN WO200261087-A2.

XX PD 08-AUG-2002.

XX PF 19-DEC-2001; 2001WO-US050107.

XX PR 19-DEC-2000; 2000US-0257144P.

XX PA (LIFE-) LIFESPAN BIOSCIENCES INC.

XX PI Burner GC, Roush CL, Brown JP;

XX PS WPI; 2003-046718/04.

XX DR P-PSDB; ABP81968.

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

XX XX

10² e

NG

No specific

Colon Cancer

QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
 Db 1861 TTTGGCAGCTTTGGCAGCATGGTGGCTGGTGGAGAATGGGGTGGTGGCCATGTCATT 1920
 QY 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu 660
 Db 1921 GGTTTTGTGCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
 QY 661 GluArgGlyPheSerValIysTyrSerAlaLysPheGluThrIysAlaProPheSerSer 680
 Db 1981 GAGCGTGGGTCTCTCTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
 QY 681 LeuIysValIleIleLeuLeuCysAlaLeuLeuLeuLeuThrMetAlaIalValProLeu 700
 Db 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
 QY 701 LeuGlyGlySerIysTyrGlyAlaSerProLeuLeuLeuLeuProLeuPheGlyGluPro 720
 Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
 QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuLeuAsnSerIleuCysPheLeuMetMet 740
 Db 2161 AGCACCATGGCTTACATGCTGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2220
 QY 741 ThrIleAlaTyrThrIysLeuTyrCysAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 760
 Db 2221 ACCATTTGCTTACACCAAGCTCTACTGCAATTTTGGCAAGGGAGACCTGGAGAAATTTGG 2280
 QY 761 AspCysSerMetValIysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
 Db 2281 GACTGCTTATGGTAAACACATTTGCCCTTGTCTTCCCAACTGCACTCTTAACCTG 2340
 QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
 Db 2341 CCTGTGCTTCTTCTGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
 QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
 Db 2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
 QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
 Db 2461 ATCTGTTCAATCTTCACTTTAAGGAGTCTGGTGAGCTGAGAAAGCAACCTACGTC 2520
 QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
 Db 2521 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTAACCTGATGTCTGAAACACAG 2580
 QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
 Db 2581 TCCTGTGACTCAACTCAAGCTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTG 2640
 QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
 Db 2641 CTCTCCAGTTCGGTCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
 QY 901 ValAlaPheValProCysLeu 907
 Db 2701 GTGGCAATTTGCCATGTCTC 2721

RESULT 3

ADC22782

ID ADC22782 standard; cDNA; 2724 BP.

XX

AC ADC22782;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human G protein-coupled receptor cDNA #39.

XX

KW Human; gene; ss; G protein-coupled receptor; GPCR;

XX

KW transmembrane-6 region; TM6; intracellular-3 region; IC3.

XX

OS Homo sapiens.
 XX US6555339-B1.
 PN 29-APR-2003.
 PD 13-OCT-1998; 98US-00170496.
 PF 14-APR-1997; 97US-00839449.
 PR 14-APR-1998; 98US-00060188.
 PR 26-JUN-1998; 98US-0090783P.
 PR 07-AUG-1998; 98US-0095677P.
 XX (AREN-) ARENA PHARM INC.
 PA Liaw CW, Behan DP, Chalmers DT;
 PI WPI: 2003-742861/70.
 DR P-PSDB; ADC22783.
 XX
 PT Creating a constitutively active version of an endogenous human G protein
 coupled receptor (GPCR) comprises substituting a specific amino acid in
 the transmembrane-6 region with a different amino acid, and testing for
 constitutive activity.
 PT
 XX Example 1; SEQ ID NO 263; 221pp; English.
 PS The invention relates to a method for treating a non-endogenous,
 CC constitutively active version of an endogenous human G protein-coupled
 CC receptor (GPCR) that has a transmembrane-6 (TM6) region and an
 CC intracellular-3 (IC3) region, by substituting a specific amino acid in
 CC the TM6 region with a different amino acid, and testing for constitutive
 CC activity. The method is useful for creating a constitutively active
 CC version of an endogenous human GPCR that comprises a transmembrane 6
 CC region and an intracellular loop 3 region. The altered human GPCR
 CC polypeptides are useful for screening test compounds for identification
 CC of inverse agonists or partial agonists of GPCR polypeptides, which may
 CC have therapeutic uses. The altered GPCRs may also be used in vivo or in
 CC vitro in biological research. A nucleic acid encoding the altered GPCR
 CC may be used to create a transgenic animal expressing the altered GPCR.
 CC The method allows screening for compounds that modulate the activity of a
 CC human G protein-coupled receptor without the need for provision of a
 CC ligand for the receptor. This is particularly useful in allowing
 CC screening of compounds against orphan receptors for which no ligand is
 CC currently known. This sequence represents cDNA encoding a human GPCR
 CC polypeptide of the invention.
 XX
 SQ Sequence 2724 BP; 665 A; 723 C; 554 G; 782 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 0 Length: 2724
 Score: 907.00 Matches: 907
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0
 US-10-751-736-84 (1-907) x ADC22782 (1-2724)
 QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
 Db 1 ATGGACACCTCCCGCTCGGTGTGCTCTGCTCTTGTGCTGTGCTGTGCTGCGGAC 60
 QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
 Db 61 GGGGGGAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCACACACTGTCTATTC 120
 QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
 Db 121 GAGCCCGACGCGCAGGATGTTGCTCAGGGTGGACTCTCCGACTCTGGGCTCTCGAGCTG 180
 QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80

Db 181 CCTTCCAACTCAGCGTCTTACCTCTACCTAGACCTCAGTATGAACAACATCAGTCAG 240
QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuLeuArgLeuAlaGlyAsn 100
Db 241 CTGCTCCCGAATCCCGTCCAGTCTCCGCTTTCCTGGAGGAGTTACGCTTTCGCGGGAAC 300
QY 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Db 301 GCTCTGACATACATTTCCCAAGGAGCATTCACCTGGCCCTTTACAGTCTTAAAGTCTTTATG 360
QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 361 CTGCAGATAATCAGCTAAGACACGTAACCCACAGAGCTCTGCAGAAATTTGGAAGCCTT 420
QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 421 CAATCCCTCGCTCGTGGTCTAACCACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 480
QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
Db 481 CTGCATTTCCCTGAGGACCTGTGGCTGGATGACAATGCGTTAACAGAAATCCCGCTCCAG 540
QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Db 541 GCTTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGGCCCTGAACAAATAACACCAC 600
QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisIleAsnAsn 220
Db 601 ATACCAGACTATGCGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATCAACAAT 660
QY 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 661 AGAATCCACTCCCTGGGAAGAAATGCTTTGTATGGGCTCCACAGCTTAGAGACTTTTAGAT 720
QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 721 TTAAATTACAATAAACCCTTGATGAATTCGCCACTGCAATTTAGGACACTCTCCAACCTTAAA 780
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 781 GAATAGGATTTTCATAGCAACAATATCAGGTGCGATACCTGAGAAAGCATTTGTAGGCAAC 840
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 841 CCTTCTCTTATTACATAATCTTCATGACAATCCCAATCCCAATTTGTTGGAGACTGCT 900
QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 901 TTTCAACATTTACCTGAACCTAAGAACACCTGACTCTGAATGGTGGCTCACAATAAATCTGAA 960
QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 961 TTTCTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1021 TCATCTCTTCTCAAACCGCTGCAATCAGTACCTTAATCTCCAAGTGTAGATCTGCT 1080
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1141 CTAAGACATAATGAATCTACGAAATTAAGTTTGAACATTTCCAGCAGTGTGCTTAGCCTC 1200
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGACAAAATTTGCTATTATTATCCACCCCAATGCATTTTCCACT 1260
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCCCTAATAAGCTGGACCTTATCGTCCAACCTCTCTGTCGTCTTTTCTATAACT 1320

QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTTAACTCATTAAATTAACAGGAATCATGCTCTACAGAGCTTGATA 1380
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTATAGAAATGCTTATGCTTACCAGTGTGT 1440
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1441 GCATTTGGAGTGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGACAAC 1500
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCATATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGACCT 1560
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1561 GACCTTTGAAGATTTCTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1621 TGTTTCACTTCCCAAGGCCCTTCAAAACCTGTGAACACCTGCTTGATGGCTGGCTGATC 1680
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1681 AGAATGGAGTGTGACCATGACGATTTCTGGACATCTTGTGTAATGCTTGTGTGACTTCA 1740
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATGGGGTCAATCGCA 1800
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGGTGGCTGGTGGTGGATCGGTTCATC 1860
QY 621 PheGlySerPheAlaArgHisGlyValaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTTTCACAGACATGCTGCTGGTGGGAGATGGGGTGGTGGATGCTCAT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu 660
Db 1921 GGTTTTTGTCTCAATTTTGTCTTTCAGAAATCTGTTTCTCTGCTTCTGCTGAGCCCTG 1980
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaPheSerSer 680
Db 1981 GAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTATATCATTTTGTCTGTGCCCTGCTGGCTTGACCATGGCCGAGTTCCTCTG 2100
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2101 CTGGGTGGCAGCAAGTATGGGCTCCCTCTCTGCTGCTTTCCTTTTGGGAGGCC 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuAsnSerIleuCysPheLeuMetMet 740
Db 2161 AGCACCATGGGGTACATGCTGCTCATCTTGTCTCAATTCCTTTGCTTCATGATG 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2221 ACCATGGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACTGGAGATATTGG 2280
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2281 GACTGCTCTATGTAACACACATTTGCCCTGTTGCTCTTCAACCACTGCATCTCAAACTGC 2340
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2341 CCTGTGGCTTTCTGTCTCTCTCTCTTTTAAATAAACCTTTACATTTTATCAGTCTGAAGTA 2400

Qy	801	IleLysPheIleLeuLeuValValProLeuProAlaCysIeuAenProLeuLeuTyr	820
Db	2401	ATTAAAGTTTATCCTTCTGGGTAGTCCCACTTCTGCATGTCTCAATCCCTTCTCTAC	2460
Qy	821	IleLeuPheAenProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal	840
Db	2461	ATCTTGTTTCAATCCTCACTTTAAAGGAGGATCTGGTGAGCGCTGAGAAAGCAAACTACGTC	2520
Qy	841	TrpThrArgSerLysHisProSerLeuMetSerIleAenSerAspAspValGluLysGln	860
Db	2521	TGCACAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTCTGATGATCGAAAAACAG	2580
Qy	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880
Db	2581	TCCTGTGACTCAACTCAAGCCTGGTAACCTTTACCAAGCTCCAGCATCATTTATGACCTG	2640
Qy	881	ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900
Db	2641	CCTCCCAAGTTCCTGGGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT	2700
Qy	901	ValAlaPheValProCysLeu	907
Db	2701	GTGGCATTTTGTCCATGCTCTC	2721

RESULT 4

ADH14255	ADH14255 standard; cDNA; 2724 BP.
XX	
XX	ADH14255;
XX	
XX	11-MAR-2004 (first entry)
DT	
XX	Human HG38 cDNA.
DE	
XX	human; non-endogenous; G protein-coupled receptor; GPCR; ss; gene.
KW	
XX	Homo sapiens.
OS	
XX	US2003105292-A1.
PN	
XX	05-JUN-2003.
PD	
XX	20-SEP-2002; 2002US-00251385.
XX	
PF	
XX	26-JUN-1998; 98US-0090783P.
PR	
XX	07-AUG-1998; 98US-0095677P.
PR	
XX	13-OCT-1998; 98US-00170496.
PR	
XX	
XX	(LIAW/) LIAW C W.
XX	(BEHA/) BEHAN D P.
PA	
XX	(CHAL/) CHALMERS D T.
XX	
XX	Liaw CW, Behan DP, Chalmers DT;
PI	
XX	WPI; 2003-801247/75.
XX	P-PSDB; ADH14256.
DR	
XX	
XX	New constitutively active, non-endogenous version of an endogenous human
PT	G protein-coupled receptor for the identification of therapeutic
PT	compounds, such as agonists.
XX	
XX	Example 1; SEQ ID NO 263; 227pp; English.
PS	
XX	
XX	The invention relates to a constitutively active, non-endogenous version
CC	of an endogenous human G protein-coupled receptor (GPCR). The GPCR is
CC	used for screening therapeutic compounds as inverse agonists, agonists or
CC	partial agonists. The GPCR can be also be used to elucidate and
CC	understand the roles of GPCRs in normal and diseased humans. The GPCR
CC	need not be purified and isolated to be used to screen for therapeutic
CC	compounds. The utility of the GPCR as a research tool is enhanced because
CC	the role of a particular receptor can be understood before the endogenous
CC	ligand is identified. The present sequence is used in the exemplification
CC	of the present invention.
CC	

301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db TTTTCAACATTTACCTTGAACTAAGACACAGTACTCTGAATGGTGGCTTCACAATAACTGAA 960
321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db TTTTCTCTGATTTAACTTGAACCTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCAAGTGTAGATCTGTCT 1080
361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db TACAACCTTATTAGAAGATTACCCAGTTTTTCAGTCTGCCAAGAGCTTCAGAAATTTGAC 1140
381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db CTAAGACACATAATGAAATCTACGAATTTAAAGTTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTATTCACCCCAATGCATTTTCCACT 1260
421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db TTGCCATCCCTAATTAAGCTGACCTATGCTCAACCTCCTGCTCTTTTCTTATTAAT 1320
441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db GGGTTACATGGTTTAACTCACCTTAAATTAACAGCAATCATGCTTACAGAGCTTGATA 1380
461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaIleThrGlnCysCys 480
Db TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAAGTCTGT 1440
481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db GCATTTGGAGTGTGAGAAATGCCCTATAGATTTCTAATCAATGAATAAAGTGCACAC 1500
501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db AGCAGTATGACGACCTTCATTAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
521 AspLeuGluAspPheLeuAspPheGluAspLeuLysAlaIleHisSerValGln 540
Db GACCTTGAAGATTTCTGCTTACATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db TGTTCACCTTCCCGAGGCCCTTCAAACTCTGAAACACCTGCTGATGGCTGGCTGATC 1680
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA 1740
581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db ACAGTTTTTCAGATGCCCTCTGACATTTCCCCCAATTTAAACTGTTAATTTGGGGTCAATCGCA 1800
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db GCAGTGAACATGCTCAAGGAGTCTCCAGTGGGTGCTGGCTGGTGGATGGATGGTTCAT 1860
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db TTTTGGCAGCTTTTGACGACATGCTGCTGGTGGAGAAATGGGGTGGTGGCCATGTCATT 1920
641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu 660
Db GGTTTTTTGTCCATTTTGTCTTCAAGAAATCATCTGTTTCTGCTTACTCTGGAGCCCTG 1980
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680

1981 GAGCGTGGTCTCTCTGTGAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db CTGAAAGTAATATTTGCTCTGTGCCCTGTGGCTTTGACCATGGCGCAGTTCCCTCG 2100
701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db CTGGGTGGCAGCAGATGATGGGCCCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGAGCC 2160
721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db AGCACCATTGGGTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCCTCATGATG 2220
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuAsnIleTyr 760
Db ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGCAAGGGGAGACTGGAGAATATTGG 2280
761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db GACTGCTCTATGGTAAACACATGCTGCTTGTCTTTCACCAACTGCATCTTAACCTGC 2340
781 ProValAlaPheLeuSerPheSerSerLeuLeuAsnLeuThrPheIleSerProGluVal 800
Db CCTGTGGCTTTCTGCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2400
801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db ATTAAGTTTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db ATCTTGTTCATCTCTCACTTTAAGAGGATCTGGTGAGCTTGAGAAAGCAAACTACGTC 2520
841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db TGGCAGACAGTCAAAACACCAAGCTTGATGTCAATTAACCTCTGATGATGTCGAAAAACAG 2580
861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2640
881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db CCTCCAGTTCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
901 ValAlaPheValProCysLeu 907
Db GTGGCAATTTGCCCATGTCTC 2721
RESULT 5
ADI32985
ID ADI32985 standard; DNA; 2724 BP.
XX AC ADI32985;
XX DT 22-APR-2004 (first entry)
XX DE Human G protein-coupled receptor (GPCR) 49 DNA.
XX KW G protein-coupled receptor 49; GPCR; neuroprotective; neural;
XX KW endocrine system disorder; gene therapy; antisense therapy; human; ds.
XX OS Homo sapiens.
XX PN US2003235910-A1.
XX PD 25-DEC-2003.
XX PF 17-JUN-2002; 2002US-00174456.
XX PR 17-JUN-2002; 2002US-00174456.

Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTtpGluAanGlyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTTGCACAGCATGGTCCCTGGTGGGAGAAATGGGTGGTGGCATGTCAAT 1920
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1921 GGTGTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTCTGCTTACTCTGGCAGCCCTG 1980
Qy 661 GluArgGlyPheSerValIysTyrSerAlaIysPheGluThrLysAlaProPheSerSer 680
Db 1981 GAGCGTGGGTTCCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 681 LeuIysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAATCATTTTGTCTCTGTGCCCTGTGGCTTGACCATGGCGCAGTTCCCTCG 2100
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2101 CTGGGTGGCAGCAGATGATGGCGCTCCCTCTCTCTGCTTGTGCTTTTGGGAGGCC 2160
Qy 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
Db 2161 AGCACCATGGGTACATGCTGCTCTCATCTTGTCTCAATTTCCCTTTGCTTCTCATGATG 2220
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2221 ACCATTGCTTACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTGG 2280
Qy 761 AspCysSerMetValIysHisIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 780
Db 2281 GACTGCTCTATGTAACACACATTGCCCTGTGCTTCTTCAACCACTGCATCTAAACTGC 2340
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTCTTAAATAAAACCTTACATTTATCAGTCTCAAGTA 2400
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2401 ATTAAGTTTAUCCCTTCTGGTGTAGTCCACCTTCTGCAATGCTCAATCCCTTCTCTAC 2460
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2461 ATCTGTTTCAATCCCTCACCTTTAAGGAGGATCTGGTGAGCCTGAGAAAGCAACCTACGTC 2520
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2521 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTAATCTCTGATGATGTGCAAAACAG 2580
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2581 TCTGTGACTCAACTCAAGCTTGGTAACTTTTACAGTCCAGCATCATCTTATGACCTG 2640
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2641 CTCTCCAGTTCCTGGTCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
Qy 901 ValAlaPheValProCysLeu 907
Db 2701 GTGGCATTTGTCCCATGTCTC 2721

RESULT 7
ADN39796
ID ADN39796 standard; cDNA; 2880 BP.
XX
AC ADN39796;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:C168.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;

KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine; gene; ss.
XX Homo sapiens.
XX WO2003042661-A2.
XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-036809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-037246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Afar P, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;
XX WPI: 2003-468649/44.
DR P-PSDB; ADN40013.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
PS Claim 8; SEQ ID NO C168; 1385pp; English.
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a nucleic acid sequence of the invention.
XX
SQ Sequence 2880 BP; 710 A; 754 C; 594 G; 822 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0 Length: 2880
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 11 Gaps: 0

QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
DB 649 ATACCAGACATATGCCCTTTGGAAACCTCTCAGCTTGGTAGTTCTATCATCTCCATACAAAT 708
QY 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
DB 709 AGAATCCACTCCCTGGGAAGAAGATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 768
QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
DB 769 TTAATTAACAATAACCTTGATGAATCCCACTGCAATAGGACACTCTCCAACCTTAAA 828
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
DB 829 GAAC TAGGATTTTCATAGCAACAATATCAGTGTGATACCTGAGAAAGCATTTGTAGCAAC 888
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
DB 889 CCTTCTCTTATACATAATCTTCTATGACATCCCATCCATTTGTTGGAGATCTGCT 948
QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
DB 949 TTTCAACATTTTAACTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAAATGAA 1008
QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
DB 1009 TTTCTCGATTTAACTGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1068
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1069 TCATCTCTTCTCAAAACCTCTGCAATCAGTTACCTTAATCTCCAGTGTAGATCTGCT 1128
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1129 TACAACCTTATAGAAGATTTACCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATGAC 1188
QY 381 LeuArgHisAsnGluIleTyrGluLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1189 CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1248
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
DB 1249 CGATGCTGAATTTGGCTTGGAAACAAAATGCTATTATTACCCCAATGCAATTTTCCACT 1308
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
DB 1309 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCGTCTTTTCTATAACT 1368
QY 441 GlyLeuHisGlyLeuThrHisLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1369 GGGTTACATGGTTAACTCACTTAAATTAACAGGAAATCATGCGCTTACAGAGCTTGATA 1428
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1429 TCATCTGAANAATTTCCAGACTCAAGGTTATAGAAATGCCATTATGCTTACCAGTGCTGT 1488
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1489 GCATTTGGAGTGTGAGAAATGCCATATAAGATTTCTAATCAATGGAATAAAGGTGACAC 1548
QY 501 SerSerMetAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1549 AGCAGTATGGAGACCTTCATAGAAGAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1608
QY 521 AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB 1609 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTTACGTGACG 1668
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
DB 1669 TGTTCACTTCCCGAGGCCCTTCAAAACCTCTGAAACACTGCTGTGATGGCTGGCTGATC 1728
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580

DB 1729 AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTTACTTTGTAATCTTTGGTGACTTCA 1788
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1789 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAATTTGGGTTCATCGCA 1848
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1849 GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGTCTGGCTGGTGTGGATGGTTCACCT 1908
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTyrGluAsnGlyValGlyCysHisValIle 640
DB 1909 TTTGGCAGCTTTTGCACGACATGGTGCCTGGTGGAGAAATGGGGTGTGGTTCGCCATGCAAT 1968
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
DB 1969 GGTTTTGTGTCATTTTGTCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 2028
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 2029 GAGCGTGGGTTCCTCTGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2088
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
DB 2089 CTGAAGATTAATCATTTTGTCTCTGTGCCCTGTGGCCTTGACCATGGCCGAGTTCCTCTG 2148
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
DB 2149 CTGGGTGGCAGCAGATGATGGCGCTCCCTCTCTGCTGCTGCTTTGGCTTTGGGAGCC 2208
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
DB 2209 AGCACCATGGGCTACATGGTGCCTCTCATCTTGTCTCAATTCCTTGTCTTCCTCATGATG 2268
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGlnIleTrp 760
DB 2269 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTGGACNAGGAGACCTGGAGAAATATTGG 2328
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
DB 2329 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCACCACTGCACTCTTAACCTGC 2388
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
DB 2389 CCTGTGGCTTTCTGTCTCTCTCTCTCTTAAATAAACCTTTACATTTATCAGTCTGAAGTA 2448
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
DB 2449 ATTAAGTTTATCTCTGTGGTAGTCCCACTTCTGCAATGTCTCAATCCCTTCTCTAC 2508
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
DB 2509 ATCTGTTCAATCTCTCACTTTAAGAGGATCTGGTGGCTGAGCTGAGAAAGCAACCTACGTC 2568
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
DB 2569 TGGCAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTCTGTATGATGTGCAAAAACAG 2628
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
DB 2629 TCTGTGATCTCAACTCAAGCCTTGGTAACCTTTTACAGCTCCAGCATCTATTATGACCTG 2688
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
DB 2689 CTTCCAGGTTCCGTGCCATCACCACTTATCCAGTGCATGAGAGCTGCCATCTTTCTCT 2748
QY 901 ValAlaPheValProCysLeu 907
DB 2749 GTGGCATTTGTCCCATGCTC 2769

RESULT 9
ADR67869

Db 1158 CTAAGACATAATGAATCTACGAATTAAGTTGACACATTTCCAGCAGTTGCTTACGCTC 1217
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1218 CGATCGCTGAATTTGGCTGGACAAATATGCTATTATTACACCAATGCAATTTCCACT 1277
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1278 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTGCTTTTCCATAAAT 1337
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1338 GGGTTACATGGTTTAACTCATTAAATTAACAGGAATCATGCCCTTACAGAGCTTGATA 1397
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1398 TCATCTGAAACTTTCCAGACTCAAGGTTATAGAAATGCCATTATGCTTACCAAGTCTGT 1457
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1458 GCATTTGGAGTGTGTGAGAATGCTATAAGATTTCTAAATCAATGGAATAAAGGTGACAAC 1517
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1518 AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1577
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1578 GACCTTGAAAGATTTCTGCTGTGACTTTGAGGAAGACCTGAAGCCCTTCATTCAAGTGCAG 1637
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1638 TGTTACCTTTCCCGACGCCCCCTTCAAACCTGTGGAACACCTGTCTGATGGCTGATC 1697
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1698 AGAATTTGGAGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTGTGGTCA 1757
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1758 ACAGTTTTCAGATCCCTCTGTACATTTCCCAATTAATGTTAAATGGGGTCATCGCA 1817
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1818 GCAGTGAACATGCTCAGCGGAGTCTCCAGTGGCTGCTGGCTGGTGGTGGTGGTGGTGGT 1877
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1878 TTTGGCAGCTTTGACAGCATGCTGCTGCTGGAGAAATGGGGTGGTGGTGGTGGTGGTGGT 1937
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 1938 GGTTTTTTGTTCATTTTGTCTTCAAGAAATCATCTGTTTTTCTGCTTACTCTGGAGCCCTG 1997
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 1998 GAGCGTGGTCTCTGTGAAATATCTGCAAAATTTGAACGAAAGCTCCATTTTCTAGC 2057
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
Db 2058 CTGAAAGTAATCATTTTGTCTGTGGCTGCTGGCTTTCAGCATGGCCGAGTTCCCTCG 2117
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2118 CTGGGTGGCAGCAGATGCGCCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2177
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2178 AGCACCATGGGTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGAG 2237
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760

Db 2238 ACCATTGCTTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGATATTTGG 2297
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2298 GACTGCTCTATGTTAAACACACATTTGCTTCTTCCCAACTGCATCTCTAAACTGC 2357
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2358 CCGTGGCTTCTTGTGCTTCTCTCTTTAAATAAACCTTTACATTTATCATGCTCCTGA 2417
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2418 ATTAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2477
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2478 ATCTTGTTCATCTCTCACCTTTAAGGAGGATCTGGTGAGCTTGAGAAAGCAACCTAC 2537
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2538 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTTAACTCTGATGATGTGCAAAACAG 2597
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2598 TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTTACAGCTCCAGCATCATCTATGACCTG 2657
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2658 CTCTCCAGTTCCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCT 2717
QY 901 ValAlaPheValProCysLeu 907
Db 2718 GTGGCATTTGTCCCATGTCTC 2738
RESULT 10
ADN39627
ID ADN39627 standard; cdna; 3032 BP.
XX
AC ADN39627;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO:A227.
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-0355250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.

QY 481 AlaPheGlyValCysGluAsnAlaThrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1641 GCATTTGGAGTGTGTCAGAAATGCTATAAGATTCTTAATCAATGGAATAAAGGTGACAAC 1700
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1701 AGCAGTATGAGCAGCCTTCATAGAAGAAGATGCTGGAAATGTTTCAGGCTCAGATGAACT 1760
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1761 GACCTTGAAGATTCTGCTGTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGCAG 1820
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1821 TGTTACCTTCCCGAGCCCTTCAAAACCTGTGAACACCTGCTTGTGCTGGCTGGCTGATC 1880
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1881 AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTGGTGAATTC 1940
QY 581 ThrValPheArgSerProLeuTrpIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1941 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAATGGGGTCATCGCA 2000
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2001 GCAGTGAACATGCTACGGGAGTCTCCAGTGGCTGCTGGCTGGTGGTGGTGGTGGTGGT 2060
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 2061 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGAGAAATGGGTTGGTGGTGGTGGTGGTGGT 2120
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 2121 GGTTTTTTGTCCATTTTGTCTTCCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 2180
QY 661 GluArgGlyPheSerValLysTrpSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2181 GAGCGTGGGTCTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2240
QY 681 LeuLysValIleLeuLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
Db 2241 CTGAAGTAATCATTTTGTCTGTGCTTCTGTGCTTCTGTGCTTCTGTGCTTCTGTGCTTCT 2300
QY 701 LeuGlyGlySerLysTrpGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2301 CTGGGTGGCAGCAGTATGGCGCTCCCTCTCTGCTGCTGCTTGGCTTTGGGGAGCCC 2360
QY 721 SerThrMetGlyTrpMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2361 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2420
QY 741 ThrIleAlaThrThrLysLeuTrpCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2421 ACCATTGCTTACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATATTGG 2480
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2481 GACTGCTCTATGGTAAAAACATTTGGCCCTGTGCTTCTTCCCACTGCACTTCAACTGC 2540
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2541 CCTGTGGCTTCTTGTGCTTCTCTCTTAAATAAACCTTACATTTATCAGTCTCCTGAAGTA 2600
QY 801 IleLysPheIleLeuValValProLeuProAlaCysLeuAsnProLeuLeuTrp 820
Db 2601 ATTAAGTTTATCTCTTGTGGTAGTCCCACTTCTGCTGCTGCTTCAATCCCTTCTCTAC 2660
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTrpVal 840
Db 2661 ATCTGTTCATCTCCTTAAAGGAGATCTGGTGGAGCTTGAGAAAGCAACCTACGTC 2720
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860

Db 2721 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTAATCTGATGTGCAAAACAG 2780
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTrpAspLeu 880
Db 2781 TCTGTGACTCAACTCAAGCCTTGGTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2840
QY 881 ProProSerSerValProSerProAlaTrpProValThrGluSerCysHisLeuSerSer 900
Db 2841 CCTCCAGTTCGGTGGCATCAGCAGCTTATCCAGTGACGTGAGAGCTGCCATCTTCTCT 2900
QY 901 ValAlaPheValProCysLeu 907
Db 2901 GTGGCATTTGTCCCATGTCTC 2921
RESULT 11
ADN39530
ID ADN39530 standard; cDNA; 3032 BP.
XX
AC ADN39530;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid, SEQ ID NO: A130.
XX
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; poriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularization syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnary; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-033464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372248P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-0397755P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Agiz N, Ginsburg WM, Gish KC, Glynn R, Hevezi P;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX
DR WPI; 2003-468649/44.
DR P-PSDB; ADN39531.
XX
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 8; SEQ ID NO A130; 1385pp; English.

XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
 CC whose expression is upregulated or downregulated in specific cancers or
 CC other diseases such as angiogenic or fibrotic disorders, and to methods
 CC of determining the presence or absence of a pathological cell in a
 CC patient by detecting a nucleic acid at least 80% identical to those of
 CC the invention or by detecting a polypeptide of the invention. The
 CC invention also relates to expression vectors and host cells comprising a
 CC nucleic acid of the invention; antibodies which specifically bind a
 CC polypeptide of the invention; use of such antibodies for drug targeting;
 CC and methods of screening for modulators of activity or expression of the
 CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
 CC antibodies and methods are useful for diagnosing, prognosing and treating
 CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
 CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
 CC neovascularisation syndromes, scarring and uterine fibroids. They may
 CC also be useful in wound healing and in contraception. The present
 CC sequence represents a nucleic acid sequence of the invention.
 XX
 SQ Sequence 3032 BP; 729 A; 816 C; 648 G; 839 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0 Length: 3032
 Score: 907.00 Matches: 907
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 11 Gaps: 0

US-10-751-736-84 (1-907) x ADN39530 (1-3032)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
 DB 201 ATGGACACCTCCCGGCTGGTGTGCTGCTCTTGCCTGTGCTGTGAGCTGGCGACC 260
 QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
 DB 261 GGGGGGAGCTCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCCCACACACTGTCTATGC 320
 QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
 DB 321 GAGCCCGACGCGGAGATGTGCTCAGGGTGGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 380
 QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
 DB 381 CCTTCCAACTCAGCGCTCTACCTCTCTACCTTAGACCTCAGTATGAACAACATCAGTCAG 440
 QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
 DB 441 CTGCTCCCGAATCCCTCGCCAGTCTCCGCTTCTGGAGGAGTACGTCTTCCGGGNAAC 500
 QY 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
 DB 501 GCTCTGACATACATCCCAAGGGAGCATTCACCTGGCCCTTTACAGTCTTAAAGTCTTATG 560
 QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
 DB 561 CTGCGAATAATACAGTAAACACGTCACCATCAGTATGTGCCCTTTCAGAAAGCTT 620
 QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
 DB 621 CAATCCCTCGCTGGATGCTAACCATCATCAGTATGTGCCCTTTCAGTGGC 680
 QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluLeuProValGln 180
 DB 681 CTGCATTCCTGAGGCACCTGTGGTGGATGACAAATGCGTTAAACAGAAATCCCGCTCCAG 740
 QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
 DB 741 GCTTTTAGAAGTTATTCGGCATTTGCAAGCCATGACCTTGGCCCTTGAACAAATATACACCAC 800
 QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAn 220

DB 801 ATACAGACTATGCGCTTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 860
 QY 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
 DB 861 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 920
 QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
 DB 921 TTAATTTACATACCTTGTATGATTTCCCACTGCAATTAGGACACTCTCCAACTTAA 980
 QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
 DB 981 GAACTAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCATTTGTAGGCAAC 1040
 QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
 DB 1041 CTTCTCTCTATTATACAAATACATTTCTATGACAATCCCAATTTGTTGGGAGATCTGCT 1100
 QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyValAsnSerGlnIleThrGlu 320
 DB 1101 TTTCAACATTTTACCTGAACCTAAGAACAACACTGACTCTGAATGGTGCCTCACAAATACTGAA 1160
 QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
 DB 1161 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1220
 QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
 DB 1221 TCATCTCTCTCAAAACCTGTGCAATCAGTTTACCTTAATCTCAAGTGTCTAGATCTGTCT 1280
 QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
 DB 1281 TACAACCTATTAGAGATTTACCCAGTTTTTCAGTCTGCCAAAGCTTCAGAAATTCAC 1340
 QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
 DB 1341 CTAGACATAATGAATCTAGAAATTAAGTTAGACACTTTCCAGCAGTTGCTTAGACCTC 1400
 QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
 DB 1401 CGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTCCACT 1460
 QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
 DB 1461 TTGCATCTCCCTAATAAAGCTGACCTATCGTCCAACTCTGCTGCTTTTCTTAACT 1520
 QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
 DB 1521 GGGTTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1580
 QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
 DB 1581 TCATCTGAAACTTTTCCAGAACTCAAGGTATTAGAAATGCCCTTATGCTTACCAGTGTGT 1640
 QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
 DB 1641 GCATTTGAGGTGTGAGAAATGCCCTATAGATTTCTAATCAATGGAATAAAGGTGACAA 1700
 QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
 DB 1701 AGCAGTAGGACGACCTTCATAAGAAAGATGTGGAATGTTTTCAGGCTCAAGATGAACGT 1760
 QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLeuAlaLeuHisSerValGln 540
 DB 1761 GACCTTTGAAGATTTCTCTGCTGACTTTGAGGAAGACCTGAAAGGCCCTTCATTCAGTGCAG 1820
 QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
 DB 1821 TGTTCACTCTCCCAAGGCCCTTCAACCCCTGTGAACACTGCTGTGATGGCTGGCTGATC 1880
 QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
 DB 1881 AGAATTTGGAGTGTGGACCATAGCAGTCTCTGGCACTTACTTGTAAATGCTTTGTGTGACTTCA 1940

Db 549 GCTCTGACATATCTCCAGGAGCATTCACCTGGCCCTTACAGCTCTTAAAGTCTTATG 608
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 609 CTGCAGATAATCAGCTAAGACACGTAACCCAGAGCTCTGCAGAAATTTGCGAAGCCTT 668
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 669 CAATCCCTGCGTCTGGATCTAACCACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 728
Qy 161 LeuHisSerLeuArgHisLeuThrLeuAspAspAsnAlaLeuThrGluIleProValGln 180
Db 729 CTGCATCTCCCTGAGGACCTGTGGCTGGATGACATGCGTTAACAGAAATCCCCGTCAG 788
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Db 789 GCTTTTAGAAGTTTATCGGCAITTCGACCATGACCTTGGCCCTGAACAAATACACCAC 848
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 849 ATACCAGACTATGCGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATACCAAT 908
Qy 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 909 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGCTCCACAGCTAGAGACTTTAGAT 968
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 969 TTTAAATTACAATAAATGATGAAATTCGCCACTGCAATTTAGGACACTCTCCAACCTTAAA 1028
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 1029 GAACTAGGATTTTCATAGCAACAAATATCAGGTGCGATACCTCGAGAAAGCATTTGTAGGCAAC 1088
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 1089 CCTTCTCTTATTACAAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGACTGCT 1148
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 1149 TTTCAACATTTACCTGAACATAAGAACACACTGACTCTGAATGGTGGCTCACAAATAACTGAA 1208
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1209 TTTCTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTTAACTGGAGCACAGATC 1268
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1269 TCATCTCTTCCCTCAAAACCGTCTGCAATCAGATTACCTTAATCTCCAGTGTAGATCTGCT 1328
Qy 361 TyrAsnLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1329 TACAACTTATTAGAAGATTTACCCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1388
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1389 CTAAGACATAATGAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGACCTC 1448
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1449 CGATCCCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTATCCACCCCAATGCATTTTCCACT 1508
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1509 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACCTCTCTGCGCTTTTCTCTATAACT 1568
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1569 GGGTTACATGGTTAACTCATTAAATTTAAACAGGAAATCATGCGCTTTACAGAGCTTGATA 1628
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480

Db 1629 TCATCTGAAAACTTTTCCAGAACTCAAGGTATAGAAATGCCTTATGCTTTACCAGTCTGT 1688
Qy 481 AlaPheGlyValCysGluAsnAlaTyrIlyLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1689 GCATTTGGAGTGTGTGAGAAATGCTTAAATCAATGAATTAAGGTGACCAAC 1748
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1749 AGCATATGGAGACCTTCATAGAAGATGCTGGAAATGTTTCAGGCTCAAGATGACGT 1808
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1809 GACCTTGAAGATTTCTGCTTGGCTTTCAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1868
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1869 TGTTCACTCTCCCAAGGCCCTTCAAAACCTGTGAACCACTGCTTGATGGCTGGGTGATC 1928
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1929 AGAATGGAGTGTGGACCATGACGATTTCTGGACATTTCTGTAAATGCTTGGTGACTTCA 1988
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1989 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGTCATCGCA 2048
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2049 GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGTGGCTGGCTGGTGGATCGGTTCACT 2108
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 2109 TTTGACGCTTTGACAGCATGCTGCTGGTGGAGAAATGGGTTGGTGGCTGATGCTATT 2168
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 2169 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 2228
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2229 GAGCGTGGTGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2288
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2289 CTGAAAGTAAATCATTTTGTCTGCTGCTGCTGGCTGGCATGGCCGACGTTCCCTG 2348
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2349 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGTGCTTTTGGGAGCC 2408
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuAsnSerLeuCysPheLeuMetMet 740
Db 2409 AGCACCATGGGTACATGCTGCTCATCTTGTGCTCAATTCCTTGTCTCTCATGATG 2468
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2469 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAGGAGAGCTGGAGATATTGG 2528
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2529 GACTGCTCTATGGTAAACACATTTGCTGCTTGTGCTTCTCACCAACTGCATCTCTAACTGC 2588
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2589 CCTGTGGCTTCTTGTGCTTCTTCTTAAACCTTTACATTTATCATGTCTGAGTGA 2648
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2649 ATTAAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2708
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2709 ATCTTGTTCATCTCCTCCTTTAAGGAGGATCTGGTGGCTTGAGAAAGCAACCTACGTC 2768

|||||
1162 TTTCAACATTTACCTGAACCTAGAACACCTGACTCTGAATGGTGCCTCACAATAACTGAA 1221
QY PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db TTTCTGATTTAACTGGAACCTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1281
QY SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db TCATCTCTTCTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1341
QY TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db TACAACCTATTAGAGATTACCCAGTTTTTCAGTCTGCCAAGAGCTTCAGAAATTTGAC 1401
QY LeuArgHisAsnGluIlePyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1461
QY ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTTACCCCAATGCAATTTCCACT 1521
QY LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db TTGCCATCCCTAATAAGCTGACACTATGCTCAACCTCCTGCTCTTTTCCATTAAC 1581
QY GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db GGGTTACATGGTTTAACTCACATTAATAATTAACAGGAAATCATGCTTTACAGAGCTTGATA 1641
QY SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db TCATCTGAAACTTTTCCAGACTCAAGGTTATAGAAATGCTTATATGCTTTACCAGTGCCTGT 1701
QY AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db GCATTTGGAGTGTGAGAAATGCCCTATAAGATTTCTAATCAATGAAATAAAGTGACAAAC 1761
QY SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db AGCAGTATGGACGACCTTCATTAAGAAAGATGCTGGAATGTTTCCAGCTCAAGATGAACCT 1821
QY AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1881
QY CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db TGTTTACCTTCCCAGGCCCTTCAAACTGTGAAACCTGCTGTGATGGCTGGCTGATC 1941
QY ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA 2001
QY ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 500
Db ACAGTTTTTCAGATTCCTCTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 2061
QY AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db GCAGTGAAATCTCTACGGAGTCTCCAGTGGCGTGTGGCTGGTGTGATGGCTTCACT 2121
QY PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db TTTGGCAGCTTTTGACGACATGGTGCCTGCTGGAGAAATGGGGTTGGTTGCCATGTCAAT 2181
QY GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuThrLeuAlaLeu 660
Db GGTTTTTTGTCTCATTTTTCCTTCAAGAAATCATCTGTTTCTCTTACTCTGGCAGCCCTG 2241
QY GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680

Db 2242 GAGCGTGGTTCCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2301
QY LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db CTGAAGTAATCATTTTGTCTGTGCCCTTGTGGCTTTGACCATGGCCGCAAGTTCCCTCG 2361
QY LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db CTGGGTGGCAGCAAGATATGGCCCTCCCTCTCTGCCCTGCCCTTTGCCCTTTGGGAGGCC 2421
QY SerThrMetGlyTyrMetValAlaLeuLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db AGCACCATGGGCTACATGGTGCCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2481
QY ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db ACCATTTGCTTACCAACGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGAAATATTGG 2541
QY AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db GACTGCTCTATGGTAACACATTTGCCCTGTGTCTTTCACCACTGCATCTCTAACTGC 2601
QY ProValAlaPheLeuSerPheSerSerLeuLeuAsnLeuThrPheIleSerProGluVal 800
Db CCTGTGGCTTTCTTGTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2661
QY IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db ATTAAGTTTATCTTCTCTGTGTAGTCCCACCTTCTGCTGCTCAATTCCTCAATCCCCTTCTCTAC 2721
QY IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db ATCTTGTTCATCTCTCACTTTAAGAGGATCTGGTGAGCCTGAGAAAGCAACCTACGTC 2781
QY TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db TGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAATCTCTGATGATGTCGAAAAACAG 2841
QY SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTACCAGCTCCAGCATCACCTATGACCTG 2901
QY ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db CCTCCAGTTTCGGTGCCATCACACGCTTATCCAGTTCAGAGCTGAGAGCTGCCATCTTCTCT 2961
QY ValAlaPheValProCysLeu 907
Db GTGGCATTTGTCCCATGTCTC 2982
RESULT 14
AA23981
ID AAX23981 standard; DNA; 3297 BP.
XX AAX23981;
XX 25-JUN-1999 (first entry)
XX Human HG38 DNA.
KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain;
KW endocrine system; skeletal muscle; spinal cord; placenta; development;
KW receptor activity modulator; ss.
OS Homo sapiens.
XX WO9915660-A1.
XX 01-APR-1999.
XX 24-SEP-1998; 98WO-05019979.
XX 24-SEP-1997; 97US-0059863P.


```
QY 741 ThrIleAlaTyrThrIleLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2221 ACCATGTGCTACACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATATTTGG 2280

QY 761 AspCysSerMetVal-LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCy 780
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2281 GACTGCTCTATGAA-AAAACACATTGCCCTGTGTCTCTCACCAACTGCATCCTAAACTG 2339

QY 780 aProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVa 800
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2340 CCCTGTGGCTTTCCTGCTCTCTCTCTTAATAAACCTTACATTTATCAGTCTCTGAAGT 2399

QY 800 ILeLeuPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTy 820
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2400 AATTAAGTTTATCCTTCTGGTGGTAGTCCCACCTTCTGCAATGCTCAATFCCCCCTTCTCTA 2459

QY 820 rIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyxVa 840
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2460 CATCTTGTTCATCTCCTCACTTTAAGAGGAGATCTGGTGAGCTGAGAAAGCAACCTACGT 2519

QY 840 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGl 860
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2520 CTGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTGATGATGTCGAAAAACA 2579

QY 860 nSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyxAspLe 880
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2580 GTCCCTGTGACTCAACTCAAGCCTTGGTAACCTTTACCAGCTCCAGCATCACTTATGACCT 2639

QY 880 uProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSe 900
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2640 GCCTCCAGTTCGGTGCCATCACCAAGCTTATCCAGTGAAGAGCTGCCATCTTTCCTC 2699

QY 900 rValAlaPheValProCysLeu 907
Db |||||||||||||||||||
2700 TGTGGCATTTGTCCCATGTCTC 2721
```

Search completed: July 12, 2005, 08:18:07
Job time : 1325 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: July 12, 2005, 07:56:07 ; Search time 2965 Seconds
(without alignments)
1920.553 Million cell updates/sec
Title: US-10-751-736-84
Perfect score: 907
Sequence: 1 MDTSLRGVLLSLFVLLQLAT.....PAYPTESCHLSLVAFPVCL 907

Scoring table: OLIGO Xgapop 60.0, Xgapext 60.0
Xgapop 60.0, Ygapext 60.0
Ygapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 6330945 seqs, 3139162390 residues

Word size: 1

Total number of hits satisfying chosen parameters: 12650801

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=xlh
-MODEL=frame+ p2n.model
-O=/cgn2_1/USPTO_spool/US10751736/runat_12072005_083643_22228/app_query.fasta_1.1095
-DB=Published Applications NA -QWTF=fast -SUFFIX=oli.Tnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MAIRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USBR=US10751736 @CGN 1.1 859 @runat_12072005_083643_22228
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:
22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	907	100.0	2724	15	US-10-251-385-263
2	907	100.0	2724	15	US-10-225-567A-421
3	907	100.0	2724	17	US-10-174-456-4
4	907	100.0	2724	20	US-10-751-736-21
5	907	100.0	2880	17	US-10-285-027-1114
6	907	100.0	2880	21	US-10-482-029-157
7	907	100.0	2880	21	US-10-651-237-48
8	907	100.0	2880	21	US-10-782-413-48
9	907	100.0	3032	17	US-10-295-027-848
10	907	100.0	3032	17	US-10-295-027-945
11	820	90.4	2724	15	US-10-251-385-277
12	806	88.9	3438	21	US-10-505-486-205
13	621	68.5	2651	17	US-10-295-027-483
14	621	68.5	2651	17	US-10-295-027-1113
15	621	68.5	2651	17	US-10-173-999-27
16	606	66.8	2082	21	US-10-851-470-3
17	232	25.6	723	15	US-10-313-542-220
18	197	21.7	1790	17	US-10-264-049-957
19	23	2.5	687	13	US-10-027-632-2009
20	23	2.5	687	13	US-10-027-632-2009
21	22	2.4	576	13	US-10-027-632-74159
22	22	2.4	576	13	US-10-027-632-299355
23	22	2.4	576	17	US-10-027-632-74159
24	22	2.4	576	17	US-10-027-632-299355
25	19	2.1	60	10	US-09-908-975-19787
26	18	2.0	2208	10	US-09-851-595-9
27	18	2.0	2208	18	US-10-664-667-9
28	18	2.0	2487	14	US-10-270-336-1
29	18	2.0	2711	10	US-09-851-595-7
30	18	2.0	2711	18	US-10-664-667-7
31	18	2.0	2901	10	US-09-851-595-3
32	18	2.0	2901	10	US-09-851-595-12
33	18	2.0	2901	18	US-10-664-667-3
34	18	2.0	2901	18	US-10-664-667-12
35	18	2.0	2988	18	US-10-302-172-723
36	18	2.0	3119	15	US-10-225-567A-580
37	18	2.0	3273	17	US-10-295-027-1119
38	18	2.0	3273	20	US-10-783-528-43
39	18	2.0	3325	17	US-10-331-496A-52
40	18	2.0	3349	14	US-10-270-336-4
41	18	2.0	3381	10	US-09-970-944-5
42	18	2.0	3438	14	US-10-176-847-89
43	18	2.0	3443	17	US-10-331-496A-94
44	18	2.0	3492	10	US-09-851-595-10
45	18	2.0	3492	18	US-10-664-667-10

ALIGNMENTS

RESULT 1
US-10-251-385-263
; Sequence 263, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13


```
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 263
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-263

Alignment Scores:
Pred. No.:      0      Length:      2724
Score:          907.00    Matches:      907
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:      100.00%  Indels:      0
Db:              15      Gaps:      0

US-10-751-736-84 (1-907) x US-10-251-385-263 (1-2724)

QY      1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
DB      1 ATGGACACCTCCCGGCTGGTGTCTCTGTCCTTGCCCTGTGCTGCTGCGAGCTGGCGACC 60

QY      21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
DB      61 GGGGGCAGCTCTCCCGAGTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCAATTGC 120

QY      41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
DB      121 GAGCCGACGCGGAGGATGTCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGGAGCTG 180

QY      61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
DB      181 CTTTCCAACTCAGCGCTCTTCACTCTCTACCTAGACCTCAGTATGACACACACTCAGTCAG 240

QY      81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
DB      241 CTGCTCCCGAATCCCTCGCCAGTCTCGCTTCTCTGGAGGAGTTACGCTTTGGCGGAAAC 300

QY      101 AlaLeuThrTyrIleProIysGlyAlaPheThrGlyLeuTyrSerLeuIysValLeuMet 120
DB      301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGCGCTTTACAGTCTTAAAGTTCTTATG 360

QY      121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
DB      361 CTGCAGATATATCAGCTAAGACACGTAACCCACAGAGGCTCTGCAGATTTGGCAGCCCTT 420

QY      141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
DB      421 CAATCCCTGCGTCTGGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480

QY      161 LeuHisSerLeuArgHisLeuThrLeuAspAspAsnAlaLeuThrGluIleProValGln 180
DB      481 CTGCATTTCCCTGAGGACCTGTGGTGGATGACATGCGTTAAACAGAAATCCCGCTCCAG 540

QY      181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnIlyIleHisHis 200
DB      541 GCTTTTAGAGTTATCGGCATTGACGCCATGACTTGGCCCTCGACAAATAACACCCAC 600

QY      201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisHisAsnAsn 220
DB      601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAAACAAT 660

QY      221 ArgIleHisSerLeuGlyIlyIysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
DB      661 AGAATCCCATCTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720

QY      241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuIys 260
DB      721 TTAATATACAAATAACTTGATGAAATTTCCCACTGCAATTAGACACTCTCCAAACCTTAAA 780

QY      261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluIyAlaPheValGlyAsn 280
DB      781 GAACTAGGATTTTCATAGCAACAATATCAGTTCGATACCTGAGAAAGCATTTGTAGGCAAC 840
```

```
QY      281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
DB      841 CCTTCTCTTATTACATACATTTCTATGACAAATCCATCCAAATTTGTTGGGAGATCTGCT 900

QY      301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
DB      901 TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGCTGCTCACAATAACTGAA 960

QY      321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
DB      961 TTTCTGCTGATTTAACTGGAACTGCACAACTCGAGAGTCTGACTTTAACTTGGAGCACAGATC 1020

QY      341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB      1021 TCATCTCTCTCTCAAAACGCTCTGCAATCAGTTACTTAATCTCCAAGTCTGATCTGCTCT 1080

QY      361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB      1081 TACAACCTATTAGAAGATTTTACCAGATTTTTCAGTCTGCCAAAGCTTCAGAAAATTGAC 1140

QY      381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB      1141 CTAAGACATAATAGAAATCTAGAAAATTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200

QY      401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
DB      1201 CGATCGCTGGAATTTGGCTTGGAAACAAAATTCCTATTATTACCCCAATGCAATTTTCCACT 1260

QY      421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
DB      1261 TTGCCATCCCTAAATAGAGCTGACCTATCGTCCAACCTCCTGCTGCTTTTCTTATAACT 1320

QY      441 GlyLeuHisGlyLeuThrHisLeuIysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB      1321 GGGTTTACATGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTTGATA 1380

QY      461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB      1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTCTGT 1440

QY      481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB      1441 GCATTTGGAGTGTGTGAGAATGCCTATAAGATTTCTTAATCAATGCAATAAAGGTGACAAC 1500

QY      501 SerSerMetAspAspLeuHisLysIysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB      1501 AGCAGTATGGACGACCTTTCAAGAAAGATCTCGAATGTTTTCAGGCTCAAGATGAACGT 1560

QY      521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB      1561 GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGAACCTGAAAAGCCCTTCATTCAGTGGCAG 1620

QY      541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleIle 560
DB      1621 TGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAACACCTGCTGTGATGGCTGGCTGATC 1680

QY      561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB      1681 AGAATTTGAGTGTGGACCATAGCAGTTCTGCACTTACTTGTAAATGCTTTGGTGAATTTCA 1740

QY      581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB      1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGGTCAATCGCA 1800

QY      601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB      1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTCCCGTGGCTGGCTGGTGGTGGATGCGTTCAT 1860

QY      621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB      1861 TTTGGCAGCTTTGACACGACATGCTGGTGGTGGGGAATGGGGTGGTGGTGGCATGTCATT 1920
```


Db 841 CTTCTCTATTACATAACATTTCTATGACATCCATCCATTTGTTGGGAGATCTGCT 900
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAenGlyAlaSerGlnIleThrGlu 320
Db 901 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTGAATGGTGGCTCACAATAACTGAA 960
Qy 321 PheProAenLeuThrGlyThrAlaAenLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACCTGGAGCAGATC 1020
Qy 341 SerSerLeuProGlnThrValCysAenGlnLeuProAenLeuGlnValLeuAenLeuSer 360
Db 1021 TCATCTCTCTCAAAACCGCTGCGAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
Qy 361 TyrAenLeuLeuGluAenLeuProSerPheSerValCysGlnLysLeuGlnLysIleAenP 380
Db 1081 TACAACTATTAGAAAGATTTACCCAGTTTTTCAGTCTGCCAAAGCTTCAGAAAAATTGAC 1140
Qy 381 LeuArgHisAenGluIleTyGluIleLysValAenThrPheGlnGlnLeuLeuSerLeu 400
Db 1141 CTAAGACATAATGAATCTACGAAATTAAGATTGACACTTTCCAGCAGTTGCTTAGCCCTC 1200
Qy 401 ArgSerLeuAenLeuAlaTriAenLysIleAlaIleIleHisProAenAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGACAAATTCGTATTATTCACCCCAATGCATTTCCACT 1260
Qy 421 LeuProSerLeuIleLysLeuAenLeuSerSerAenLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCCCTAAATAAGCTGACCTATGCTCAACCTCCTGCTCTTTTCCCTATAACT 1320
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAenHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTTAACTACACTTAAATTAACAGGAAATCATGCTTCACAGAGCTTGATA 1380
Qy 461 SerSerGluAenPheProGluLeuLysValIleGluMetProTyAlaTyGlnCysCys 480
Db 1381 TCATCTGAATACTTTCCAGAACTCAAGGTATAGAAATGCCCTTATGCTTACCAGTCTGT 1440
Qy 481 AlaPheGlyValCysGluAenAlaTyLysIleSerAenGlnTrpAenLysGlyAenPasn 500
Db 1441 GCATTTGGAGTGTGTGAGAATGCCCTATAGATTTTCTAATCAATGAATAAAGGTGACAA 1500
Qy 501 SerSerMetAenAenLeuHisLysIleAspAlaGlyMetPheGlnAlaGlnAenGluArg 520
Db 1501 AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Qy 521 AspLeuGluAenPheLeuLeuAenPheGluGluAenLeuLysAlaLeuHisSerValGln 540
Db 1561 GACCTTGAAGATTCTCTGCTTGACTTTGAGGAAGACCTGMAAGCCCTTCATCTCAGTGCAG 1620
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAenGlyTriPheIle 560
Db 1621 TGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAACACCTGCTTGATGGTGGCTGATC 1680
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAenAlaLeuValThrSer 580
Db 1681 AGAATTGGAGTGTGGACCATAGCAGTCTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCA 1740
Qy 581 ThrValPheArgSerProLeuTyriIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCATTAACCTGTTAAATGGGGTCAATCGCA 1800
Qy 601 AlaValAenMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAenAlaPheThr 620
Db 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGT 1860
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAenGlyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTTGGACACGATGGTGGTGGTGGGAGATGGGGTGGTGGTGGTGGTGGTGGT 1920
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660

Db 1921 GGTTTTTTGTCCATTTTTTGTCTTCAGAAATCATCTGTTTTTCTGCTCTTACTCTGGCAGCCCTG 1980
Qy 661 GluArgGlyPheSerValLysTyriSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 1981 GAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAATCATTTTTGTCTCTGTGCCCTGTGGCTTGACCATGGCCGCGAGTTCCCTG 2100
Qy 701 LeuGlyGlySerLysTyriGlyAlaSerProLeuCysLysLeuProLeuProPheGlyGluPro 720
Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCGCTGCTTTGGCTTTTGGGGAGGCC 2160
Qy 721 SerThrMetGlyTyriMetValAlaLeuIleLeuAenSerLeuCysPheLeuMetMet 740
Db 2161 AGCACCATGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2220
Qy 741 ThrIleAlaTyriThrLysLeuTyriCysAenLeuAenLysGlyAenLeuGluAenIleTrp 760
Db 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATTTGG 2280
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAenCysIleLeuAenCys 780
Db 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCCCAACTGCATCTCTAAACTGC 2340
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAenLeuThrPheIleSerProGluVal 800
Db 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Qy 801 IleLysPheIleLeuLeuValValValProLeuProAlaCysLeuAenProLeuLeuTyri 820
Db 2401 ATTAAGTTTATCTCTTCTGTGGTGGTAGTCCACTTCTGCAATGCTCAATCTCTCTCTAC 2460
Qy 821 IleLeuPheAenProHisPheLysGluAenLeuValSerLeuArgLysGlnThrTyriVal 840
Db 2461 ATCTTGTTCATCTCACTTTTAAAGGAGATCTGTGAGCCTGAGAAACCAACTACGTC 2520
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAenSerAenAenValGluLysGln 860
Db 2521 TGGCAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTCTGATGATGTGCAAAAACAG 2580
Qy 861 SerCysAenSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyriAspLeu 880
Db 2581 TCCGTGACTCAACTCAAGCCTTGTAACCTTTACAGCTCCAGCATCATTATGACCTG 2640
Qy 881 ProProSerSerValProSerProAlaTyriProValThrGluSerCysHisLeuSerSer 900
Db 2641 CCTCCAGTTCCGTGCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCTCT 2700
Qy 901 ValAlaPheValProCysLeu 907
Db 2701 GTGGCATTTGTCCCATGTCTC 2721

RESULT 3

US-10-174-456-4

; Sequence 4, Application US/10174456

; Publication No. US20030235910A1

; GENERAL INFORMATION:

; APPLICANT: Susan M. Preiser

; TITLE OF INVENTION: ANTISENSE MODULATION OF G PROTEIN-COUPLED RECEPTOR 49 EXPRESSION

; FILE REFERENCE: RIG-0374

; CURRENT APPLICATION NUMBER: US/10/174,456

; NUMBER OF SEQ ID NOS: 139

; SEQ ID NO 4

; LENGTH: 2724

; TYPE: DNA

; ORGANISM: H. sapiens

; FEATURE:

; NAME/KEY: CDS

Db 901 TTTCAACATTTTACCTGAACACACTGCTGTAATGGTGGCTCAAAATACAGAA 960
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyValaGlnIle 340
Db 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGCTGACCTTTAACTGGAGCACAGATC 1020
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1021 TCATCTCTTCTCAAAACCGTTCGAATCAGTTACCTAATCTCCAAGTGTAGATCTGTCT 1080
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLeuLeuGlnIleAsp 380
Db 1081 TACACACCTATTAGAGATTTACCCAGTTTTCAGTCTGCCAAAGACTTCAGAAAAATTGAC 1140
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1141 CTAAGACATAATGAATCTACGAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT 1260
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCTCTTTTCTATAACT 1320
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTTAACTCATTAAATTAACAGGAATCATGCTTCCAGAGCTTGATA 1380
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTACCAGTGTGT 1440
Qy 481 AlaPheGlyValCysGluAsnAlaTyrIlylSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1441 GCATTTGGAGTGTGTGAGAATGCCCTATAGAAATTTCTTAATCAATGAATAAAGGTGACAA 1500
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCAGTATGGAGCCTTCATAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGACGT 1560
Qy 521 AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1561 GACCTTGAAGATTTCTGCTGACTTTTCAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleuIle 560
Db 1621 TGTTCCACCTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGCTTGTATGGCTGGCTGATC 1680
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1681 AGAATGGAGTGTGGACCATAGCAGTCTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCA 1740
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCAATTAACCTGTTAAATGGGGTCAATGCA 1800
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1860
Qy 621 PheGlySerPheAlaArgHisGlyValaTrpTrpGluAenglyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTTGGACAGCATGGTGGCTGGGGAGAAATGGGTTGGTGGCATGTGCAAT 1920
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1921 GGTTTTTTGTCCATTTTGTCTCAGAACTCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680

Db 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAAATCATTTTGTCTGTGGCTCTGTGGCTTGGACCATGGCGCAGTTCCCTCG 2100
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2101 CTGGTGGCAGCAGATATGGGCGCTCTCCCTCTCTGCTGCTGCTTTTGGCGAGCC 2160
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2161 AGCACCATGGGCTACATGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAATATTGG 2280
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2281 GACTGCTCTATGTGTAACACATTTGCTCTTCCCAACTGCATCTCTAAACTGC 2340
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2341 CCTGTGGCTTTCTTGTCTCTTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2401 ATTAAGTTTATCTTCTGTGTGTAGTCCACTTCTCTGATGCTCAATCCCTCTCTTAC 2460
Qy 821 IleLeuPheAsnProHisPheLysGlyAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2461 ATCTTGTTCATCTCTCACTTTAAGGAGGATCTGGTGAGCCTGAGAAAGCAACCTACGTC 2520
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2521 TGGCAAGATCAAAACACCCCAAGCTTGAATCAATTAATCTGTATGATGTGAAAAACAG 2580
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2581 TCCTGTGACTCACTCAAGCCTTGGTAACCTTTACCAGCTCCAGCATCATTATGACCTG 2640
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2641 CTTCCAGTTCCTGTGCCATCACCAGCTTATCCAGTACTGAGAGCTGCCATCTTCTCTCT 2700
Qy 901 ValAlaPheValProCysLeu 907
Db 2701 GTGGCATTTGTCCCATGTCTC 2721

RESULT 5

US-10-295-027-1114

Sequence 1114, Application US/10295027

Publication No. US2003023350A1

GENERAL INFORMATION:

APPLICANT: ~~Asa~~XDaniel

APPLICANT: Aziz, Natasha

APPLICANT: Ginsberg, Wendy M.

APPLICANT: Gish, Kurt C.

APPLICANT: Glynn, Richard

APPLICANT: Hevezi, Peter A.

APPLICANT: Mack, David H.

APPLICANT: Murray, Richard

APPLICANT: Watson, Susan R.

APPLICANT: Eos Biotechnology, Inc.

TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

TITLE OF INVENTION: Methods of Screening for Modulators of Cancer

FILE REFERENCE: 018501-012500US

CURRENT APPLICATION NUMBER: US/10/295,027

CURRENT FILING DATE: 2002-11-13

PRIOR APPLICATION NUMBER: US 09/663,733

PRIOR FILING DATE: 2000-09-15

PRIOR APPLICATION NUMBER: US 60/350,666

1020
seg 1380
coln
par 68


```
Db 1669 TGTTCACCTTCCCGAGCCCTTCAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC 1728
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1729 AGAAATGGAGTGTGGACCATAGCAGTCTGGCACCTTACTTGTAAATGTTGGTGAATTCA 1788
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleIysLeuLeuIleGlyValIleAla 600
Db 1789 ACAGTITTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTAAATGGGGTCAATCGCA 1848
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1849 SCAGTGAACATGCTCAGGGAGTCTCAGTGCCTGCTGGCTGGTGTGGATGGTTCAT 1908
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1909 TTTGGCAGCTTTTGACAGACATGCTGCTGTGGAGAAATGGGGTGGTGGCCATGTCATT 1968
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheIleLeuThrLeuAlaLeu 660
Db 1969 GGTITTTTGTTCATTTTGTCTTGAATCATCTGTCTTCTGCTTACTCTGGAGCCCTG 2028
Qy 661 GluArgGlyPheSerValIysTyrSerAlaIysPheGluThrIysAlaProPheSerSer 680
Db 2029 GAGCGTGGGTCTCTGTGAATATCTGCAAAATTTGAACGAAAGCTCCATTTTCTAGC 2088
Qy 681 LeuIysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2089 CTGAAGTAAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148
Qy 701 LeuGlyIysSerIysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2149 CTGGGTGGCAGCAAGTATGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
Qy 721 SerThrMetGlyTyrMetValAlaLeuLeuIleLeuLeuAsnSerLeuCysPheLeuMet 740
Db 2209 AGCACCATGGGTACATGCTGCTGCTCACTTGTCTCAATTCCTTTGCTTCTCTCATGAG 2268
Qy 741 ThrIleAlaTyrThrIysLeuTyrCysAsnLeuAspIysGlyAspLeuGluAsnIleTrp 760
Db 2269 ACCATTGCTTACACCAAGCTCTACTGCAATTTGGCAAGGAGACCTGGAGAAATATTGG 2328
Qy 761 AspCysSerMetValIysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2329 GACTGCTCTATGTTAAACACATTTGCTGCTTCTTCCCAACTGCATCTCTAAACTGC 2388
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2389 CCTGTGGCTTTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2448
Qy 801 IleIysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2449 ATTAAGTTTATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2508
Qy 821 IleLeuPheAsnProHisPheIysGluAspLeuValSerLeuArgIysGlnThrTyrVal 840
Db 2509 ATCTTGTTCATCTCTTAAAGAGGATCTGGTGGAGCTCTGAGCTCGAGAAAGCAAACTAC 2568
Qy 841 TrpThrArgSerIysHisProSerLeuMetSerIleAsnSerAspValGluIysGln 860
Db 2569 TGGACAAGATCAAAACCAACCAAGCTTGATGTCAATTAATCTGTATGATGTGCAAAACAG 2628
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2629 TCTGTGATCTCACTCAAGCCTTGGTAACTTTTACAGGTCCAGCATCACTTATGACCTG 2688
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2689 CTTCCAGTTCCTGGTGCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2748
Qy 901 ValAlaPheValProCysLeu 907
Db 2749 GTGGCATTTGTCCCATGTCTC 2769
```

RESULT 6

```
US-10-482-029-157
; Sequence 157, Application US/10482029
; Publication No. US20050037445A1
; GENERAL INFORMATION:
; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482,029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
; LENGTH: 2880
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-482-029-157
```

Alignment Scores:
Pred. No.: 0 Length: 2880
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-10-751-736-84 (1-907) x US-10-482-029-157 (1-2880)

```
Qy 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
Db 49 ATGGACACCTCCCGGCTCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 108
Qy 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Db 109 GGGGGCAGCTCTCCAGGCTCTGGTGTGCTGCTGAGGGGCTGCCCCACACACTGTCTATG 168
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 169 GAGCCGCGAGCGGAGGATGTTGCTCAGGGTGGAGTCTCCGACCTGGGGCTCTCGAGCTG 228
Qy 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Db 229 CCTTCCACCTCAGCGTCTTACCTCTCTACCTAGACCTCAGTATGAACACATCAGTCAG 288
Qy 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
Db 289 CTGCTCCCGAATCCCTGCGGCTCTCCGCTTCTGGAGAGTTAGCTCTTGGGGAAAC 348
Qy 101 AlaLeuThrTyrIleProIysGlyAlaPheThrGlyLeuTyrSerLeuIysValLeuMet 120
Db 349 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGGCTTTTACAGTCTTAAAGTTCTATG 408
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 409 CTGCAAGATAATACAGCTAAGACACAGTACCACAGAGCTCTCAGAAATTTGGAGACCTT 468
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 469 CAATCCCTGCGTCTGGATGCTTAACCATCATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 528
Qy 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluLeuProValGln 180
Db 529 CTGCATTTCCCTGAGGACCTGTGGCTGGATGACATGCGTTAACAGAAATCCCGCTCCAG 588
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaIleAsnIysIleHis 200
Db 589 GCTTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGGCCCTTGAACAAAAATACACCAC 648
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsn 220
Db 649 ATACCAGACTATGCTCTTTGGAAACCTTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 708
```

AGD

Qy	621	PheGlySerPheAlaArgHisGlyValAlaTrpTrpGluAsnGlyValGlyCysHisValIle	640
Db	1909	TTTGCCAGCTTTGGACGACATGGTGCCTGGTGGGAGAAATGGGGTTGGTGGCCATGTCTATT	1968
Qy	641	GlyPheLeuSerIlePheAlaSerGluSerSerValPheIleuThrIleuAlaAlaLeu	660
Db	1969	GGTTTTTTGGTCCATTTTTGGCTTCAGAAATCATCTGTTTTCTCGCTTACTCTGGCAGGCCCTG	2028
Qy	661	GluArgGlyPheSerValIleTyrSerAlaIlePheGluThrIleValAlaProPheSerSer	680
Db	2029	GAGCGTGGGTCTCTGTGAAATATTTGCAAAATTTGAACGAAAGCTCCATTTTCTAGC	2088
Qy	681	LeuIysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu	700
Db	2089	CTGAAAGTAATCATTTTTGCTCTGTGCCCTGTGGCCCTTGACCATTGGCGCGCAGTTCCGCTG	2148
Qy	701	LeuCllyGlySerIysTyrGlyAlaSerProIleuCysIleuProIleuProPheGlyGluPro	720
Db	2149	CTGGGTGGCAGCAAGTAGTGGCGCTCCCTCTCTCGCTGCTTTGGCTTTTGGGGAGCCC	2208
Qy	721	SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet	740
Db	2209	AGCACCATGGCTACATGGTGGCTCTCATCTTGCTCAATTTCCCTTTGCTTCTCATGATG	2268
Qy	741	ThrIleAlaTyrThrIleValLeuTyrCysAsnLeuAspIysGlyAspLeuGluAsnIleTrp	760
Db	2269	ACCATTGCCCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGGAGAATATTGG	2328
Qy	761	AspCysSerMetValIleValHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys	780
Db	2329	GACTGCTCTATGGTAAACACATTTGCCCTGTGTGCTCTTCACCAACTGCATCTCAAACTGC	2388
Qy	781	ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal	800
Db	2389	CCTGTGGCTTTCTTGTCCTCTCTCTTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA	2448
Qy	801	IleIysPheIleLeuLeuValValProIleuProAlaCysIleuAsnProLeuLeuTyr	820
Db	2449	ATTAAGTTTATCTCTTCTGGTGGTAGTCCCACTTCTCTGCATGTCTCAATCCCTTCTCTAC	2508
Qy	821	IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgIysGlnThrTyrVal	840
Db	2509	ATCTTGTTCATCTCTACTTTAAGGAGGACTGGTGGAGCTTGAAAGCAAACTACGTC	2568
Qy	841	TrpThrArgSerIysHisProSerLeuMetSerIleAsnSerAspValGluLysGln	860
Db	2569	TGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTCTGATGATGTCGAAAAACAG	2628
Qy	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880
Db	2629	TCCTGTGACTCAACTCAAGCGCTTGGTAACCTTTACCAGCTCCAGACTCATTTATGACCTG	2688
Qy	881	ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900
Db	2689	CCTCCAGTTCGGTGCCATCACCACTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT	2748
Qy	901	ValAlaPheValProCysLeu	907
Db	2749	GTGGCATTTGTCCCATGTCTC	2769

```

; PRIOR FILING DATE: 2003-08-28
;
; NUMBER OF SEQ ID NOS: 94
;
; SOFTWARE: PatentIn version 3.1
;
; SEQ ID NO 48
;
; LENGTH: 2880
;
; TYPE: DNA
;
; ORGANISM: human
;
US-10-782-413-48

```

Alignment Scores:			
Pred. No.:	0	Length:	2880
Score:	907.00	Matches:	907
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-10-751-736-84 (1-907) x US-10-782-413-48 (1-2880)

Qy	1	Meta	Thr	Ser	Arg	Leu	Gly	Val	Leu	Leu	Ser	Leu	Pro	Val	Leu	Leu	Gln	Leu	Ala	Thr	20	
Db	49	ATG	GAC	ACC	TCC	GCG	CTC	GGT	GCT	CTC	TGT	CTT	TGC	TGT	GCT	GCT	GCT	GCT	GCT	GAC	108	
Qy	21	Gly	Cys	Ser	Pro	Arg	Ser	Gly	Val	Leu	Leu	Arg	Gly	Cys	Pro	Thr	His	Cys	His	Cys	40	
Db	109	GGG	GCG	GAG	CTCTCC	CCAG	GTC	TGG	TGT	TGT	CTG	TAG	GGG	GCT	CC	CC	CC	CC	CC	CC	168	
Qy	41	Glut	Pro	Asp	Gly	Arg	Met	Leu	Leu	Arg	Val	Asp	Cys	Ser	Asp	Leu	Gly	Val	Ser	Leu	60	
Db	169	GAG	CCC	GAG	CGG	CAG	GAT	TG	TGCT	CAG	GGT	GG	AGCT	GCT	CCG	ACC	TGG	GGG	CTCTCG	GAG	228	
Qy	61	Pro	Ser	Asn	Leu	Ser	Val	Phe	Thr	Ser	Tyr	Leu	Asp	Leu	Ser	Met	Asn	Asn	Ile	Ser	80	
Db	229	CCT	TCCA	ACCT	CAG	CGT	CTT	CAC	CTCCT	ACCT	AG	ACCT	CTCAG	TAC	CTCAG	TAT	GAA	CAACAT	CAGT	CAG	288	
Qy	81	Leu	Leu	Pro	Leu	Pro	Ser	Leu	Arg	Phe	Ile	Leu	Glu	Leu	Arg	Leu	Ala	Gly	Asn	100		
Db	289	CTG	CTCC	GGA	TCC	CCCT	GCC	CAG	TC	CGC	CTC	TGG	GAG	GAG	TAC	CGT	TTG	CGG	GGA	AAC	348	
Qy	101	Ala	Leu	Thr	Tyr	Ile	Pro	Lys	Gly	Val	Ala	Phe	Thr	Gly	Ile	Tyr	Ser	Leu	Lys	Val	120	
Db	349	GCT	CTG	ATC	ATAC	TATCC	CA	GGG	AGC	AT	CACT	GGC	CTTT	AC	AGT	CTT	AA	AGT	CTT	TA	408	
Qy	121	Leu	Gln	Asn	Asn	Gln	Leu	Arg	His	Val	Pro	Thr	Glu	Ala	Leu	Gln	Asn	Leu	Arg	Ser	140	
Db	409	CTG	CAGA	TAAT	CAG	CTA	AG	CA	CGT	TA	ACC	CA	CGA	AGCT	CTG	CA	AGAT	TTG	CGA	AGC	468	
Qy	141	Gln	Ser	Leu	Arg	Leu	Asp	Ala	Asn	His	Ile	Ser	Tyr	Val	Pro	Pro	Ser	Cys	Phe	Ser	160	
Db	469	CAAT	CCCT	GGCT	CTG	GAT	GTAA	CCAC	AT	CAG	CTAT	GTG	CCCC	CA	AGCT	GT	TTT	CA	G	TGC	528	
Qy	161	Leu	His	Ser	Leu	Arg	His	Ile	Tyr	Trp	Leu	Asp	Asp	Asn	Ala	Leu	Thr	Glu	Ile	Pro	180	
Db	529	CTG	CAT	TTCCCT	GAG	CACT	TG	TGCT	GG	AT	GACA	CAAT	GCG	TTAA	CAG	AAAT	CCCC	CT	CC	G	588	
Qy	181	Ala	Phe	Arg	Ser	Leu	Ser	Ala	Leu	Gln	Ala	Met	Thr	Leu	Ala	Leu	Asn	Lys	Ile	His	200	
Db	589	GCT	TTT	AGA	AGT	TTAT	TGG	CAT	TGC	AG	CCAT	GAC	CTT	GCG	CCCT	GAA	CAAA	TAAT	CAC	CA	648	
Qy	201	Ile	Pro	Asp	Tyr	Ala	Phe	Gly	Asn	Leu	Ser	Ser	Leu	Val	Val	Leu	His	Leu	His	Asn	220	
Db	649	ATAC	CAG	ATAT	GCT	TTT	TGG	AAAC	CTCT	CC	AGCT	TGGT	TAGT	TTCT	TAC	ATCT	CC	ATA	CAAT	708		
Qy	221	Arg	Ile	His	Ser	Leu	Gly	Lys	Lys	Cys	Phe	Asp	Gly	Leu	His	Ile	Ser	Leu	Glu	Thr	240	
Db	709	AGA	TCC	ACAT	CCCT	CGG	AAA	AAA	ATG	CTT	GAT	GGCT	TCC	CA	CGCT	TAG	AGACT	TT	TAG	AT	768	
Qy	241	Leu	Asn	Tyr	Asn	Asn	Leu	Asp	Glu	Phe	Pro	Thr	Ala	Ile	Arg	Thr	Leu	Ser	Asn	Leu	260	
Db	769	TTAA	TTT	TACA	ATA	TA	CCCT	TGAT	GA	AT	TTCC	CACT	GC	AT	TAG	CA	CACT	CT	CC	AA	828	
Qy	261	Glu	Leu	Gly	Phe	His	Ser	Asn	Asn	Ile	Arg	Ser	Ile	Pro	Glu	Lys	Ala	Phe	Val	Gly	Asn	280

829	GAAC	TAGAGT	TTCTAT	GACAA	CAATAT	CAGG	TCGAT	CAC	TGAGAA	AGCA	TATTTG	TAGG	CAAC	888													
281	ProSer	Leu	Ile	Thr	Ile	His	Phe	Thr	Asp	Asn	Pro	Ile	Gln	Phe	Val	Gly	Arg	Ser	Ala	300							
889	CTT	TCT	CTT	ATT	ACA	ATAC	ATTT	CT	AT	GACA	CAAT	CCCAT	CCAA	TTT	TG	TGG	GAG	AT	CT	948							
301	Phe	Gln	His	Ile	Ser	Leu	Pro	Glu	Leu	Arg	Thr	Ile	Thr	Leu	Asn	Gly	Ala	Ser	Gln	Ile	Thr	Glu	320				
949	TTT	CAA	CA	TTT	TAC	CT	GAA	CT	TAAG	AACA	CT	TG	ACT	CT	GAA	TG	TG	CT	CAC	AAA	TAA	CT	GAA	1008			
321	Phe	Pro	Asp	Leu	Thr	Gly	Thr	Ala	Asn	Leu	Glu	Ser	Leu	Thr	Leu	Thr	Gly	Ala	Gln	Ile			340				
1009	TTT	CCT	GAT	TTT	AA	CT	TG	GA	CT	TG	CAAA	CT	TG	GAG	AG	TC	TG	ACT	TT	AA	CT	TG	GAG	CA	GAT	1068	
341	Ser	Ser	Leu	Pro	Gln	Thr	Val	Cys	Asn	Gln	Leu	Pro	Asn	Leu	Gln	Val	Leu	Asp	Leu	Ser			360				
1069	TCAT	CT	CT	CT	CAA	CG	CT	CA	AT	CAG	TTAC	CT	TA	AT	CT	CA	AG	TG	CT	AG	AT	CT	CT	CT	CT	1128	
361	Tyr	Asn	Leu	Leu	Glu	Asp	Leu	Pro	Ser	Phe	Ser	Val	Cys	Gln	Lys	Leu	Gln	Lys	Ile	Asp			380				
1129	TACA	AC	CT	TAT	TAGA	AG	AT	TTT	ACC	CG	AT	TTT	T	CAG	CT	CT	CC	AAA	AG	CT	T	C	G	AAA	AT	TG	1188
381	Leu	Arg	His	Asn	Glu	Ile	Tyr	Glu	Ile	Lys	Val	Asp	Thr	Phe	Gln	Gln	Leu	Leu	Ser	Leu			400				
1189	CTA	GAC	AT	AT	GAA	AT	CT	CA	AA	TTAA	AG	T	GAC	ACT	TT	CC	AG	AG	T	TG	CT	T	G	CC	T	1248	
401	Arg	Ser	Leu	Asn	Leu	Ala	T	Pro	Asn	Lys	Ile	Ala	Ile	His	Pro	Asn	Ala	Phe	Ser	Thr			420				
1249	CGAT	CG	CT	GAA	TTT	GG	CT	TG	GA	CA	AAA	AT	TG	CT	AT	TAT	T	CAC	CC	CA	AT	T	G	CA	T	1308	
421	Leu	Pro	Ser	Ile	Leu	Lys	Leu	Asp	Lys	Ser	Ser	Asn	Leu	Leu	Ser	Ser	Phe	Pro	Ile	Thr			440				
1309	TTG	CA	T	CC	AT	CC	CT	TA	TAA	AG	CT	GG	CA	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	1368	
441	Gly	Leu	His	SGly	Leu	Thr	His	Leu	Lys	Leu	Thr	Gly	Asn	His	Ala	Leu	Gln	Ser	Leu	Ile			460				
1369	GGG	TTA	CAT	GG	TTT	AA	CT	CA	CT	TAAA	AT	TAA	CAG	AA	AT	CAT	GC	CT	TAC	AG	AG	CT	T	G	ATA	1428	
461	Ser	Ser	Glu	Asn	Phe	Pro	Glu	Leu	Lys	Val	Ile	Glu	Met	Pro	Tyr	Ala	Tyr	Gln	Cys	Cys			480				
1429	TCAT	CT	GAA	AA	CT	TT	CC	AG	AA	CT	CA	AG	TTAT	AG	AA	TG	CT	TAT	GT	TAC	CA	G	T	G	T	1488	
481	Ala	Phe	Gly	Val	Cys	Glu	Asn	Ala	Tyr	Lys	Ile	Ser	Asn	Gln	Trp	Asn	Lys	Gly	Asp	Asn			500				
1489	GCA	TTT	GG	AG	TG	TG	AGA	TG	CC	TA	AGA	TTT	CT	AA	TG	CA	TAA	TG	CA	AT	AA	AG	GT	G	CA	CA	1548
501	Ser	Ser	Met	Asp	Asp	Leu	His	Lys	Lys	Asp	Ala	Gly	Met	Phe	Gln	Ala	Gln	Asp	Glu	Arg			520				
1549	AGC	AG	TAT	G	AG	CA	CT	TTCA	TAA	GAA	AG	TG	CTG	GA	AT	TG	TT	C	AG	G	CT	C	AA	G	AT	G	1608
521	Asp	Leu	Glu	Asp	Phe	Leu	Leu	Asp	Phe	Glu	Glu	Asp	Leu														

Qy	641	GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu	660
Db	1969	GGTTTTTTGTCCATTTTTTGGTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	2028
Qy	661	GluArgGlyPheSerValIleValThrSerAlaIlePheGluThrIleValAlaProPheSerSer	680
Db	2029	GAGCGTGGTTTCTCTGTGAAANATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTTAC	2088
Qy	681	LeuIysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaIaValProLeu	700
Db	2089	CTGAAAGTAATCATTTTGGCTCTGTGGCCCTGCTGGCCCTTGACCATGGCGCAGTTCCCTCG	2148
Qy	701	LeuGlyGlySerIleValIleValSerProIleuCysLeuProIleuProPheGlyGluPro	720
Db	2149	CTGGGTGGCAGCAAGTAGTAGGGCCCTCCCTCTCTGCGCTGCTTTTGGCTTTTGGGGAGCCC	2208
Qy	721	SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet	740
Db	2209	AGCACCATGGCTACATGGTGGCTCTCATCTTGCTCAATTCCTTTGCTTCTCATGATG	2268
Qy	741	ThrIleAlaTyrThrIysLeuTyrCysAsnLeuAspIysGlyAspLeuGluAsnIleTyr	760
Db	2269	ACCATTGCCTACACCAAGCTCTACTGCAATTTTGACCAAGGAGACCTGGAGAAATATTTGG	2328
Qy	761	AspCysSerMetValIysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys	780
Db	2329	GACTGCTCTATGGTAAACACATGGCCCTGTGCTCTTCCACCAACTGCATCTCTAAACTGC	2388
Qy	781	ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal	800
Db	2389	CCTGTGGCTTCTTGTCCTTCTCCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA	2448
Qy	801	IleIysPheIleLeuLeuValValProIleuProAlaCysLeuAsnProLeuLeuTyr	820
Db	2449	ATTAAAGTTTATCCTTCTGTGGTAGTCCCACTTCTCTGCATGTCTCAATCCCTTCTCTAC	2508
Qy	821	IleLeuPheAsnProHisPheIysGluAspLeuValSerLeuArgIysGlnThrTyrVal	840
Db	2509	ATCTGTTCATCTCTCACTTTAGGAGGATCTGGTGGAGCTTGAGAAAGCAACCTACGTC	2568
Qy	841	TrpThrArgSerIysHisProSerLeuMetSerIleAsnSerAspAspValGluIysGln	860
Db	2569	TGGACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTCATGATGTCGAAAAACAG	2628
Qy	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880
Db	2629	TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTACCAGCTCCAGACATCATTTATGACCTG	2688
Qy	881	ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900
Db	2689	CCTCCAGTTCGGTGGCCATCACCAAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT	2748
Qy	901	ValAlaPheValProCysLeu	907
Db	2749	GTGGCATTTGTCCCATGTCTC	2769

RESIN, T. 9

US-10-295-027-848

03 10 255 027 010
; Sequence 848, Application US/10295027

; Publication No. US20030232350A1

GENERAL INFORMATION:

APPLICANT: Afar, Daniel

APPLICANT: Aziz, Natasha

APPLICANT: Ginsberg, Wendy M.

APPLICANT: Gish, Kurt

APPLICANT: Glynne, Richard

APPLICANT: Hevezi, Peter A.

APPLICANT: Mack, David H.

APPLICANT: Murray, Richard

APPLICANT: Watson, Susan R.

APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and

Db 261 GGGGCGAGCTCTCCAGGCTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCTATTGC 320
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 321 GAGCCGACGGCAGGATGTGCTCAGGGTGAGCTGCTCCGACCTGGGGCTCTCGAGCTG 380
Qy 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Db 381 CCTTCCAACTCAGCGTCTTCACTCTCCCTAGACCTCAGTATGACACATCAGTCAG 440
Qy 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
Db 441 CTGCTCCCGAATCCCTCCAGCTCTCCGCTTCTCTGGAGGATTACGTCCTTGGCGGAAC 500
Qy 101 AlaLeuThrTyrIleProGlyGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Db 501 GCTCTGACATCATTTCCCAAGGAGCATTCACCTGGCCCTTACAGTCTTTAAAGTCTTATG 560
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 561 CTGCAGAAATATCAGCTAAGACACAGTACCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 620
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 621 CAATCCCTGCGTCTGGATGCTAAACCACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 680
Qy 161 LeuHisSerLeuArgHisLeuThrLeuAspAspAsnAlaLeuThrGluIleProValGln 180
Db 681 CTGCATTCCTCGAGCACCTGGCTGGTGGATGACATGCGTTTAAACAGAAATCCCGTCCAG 740
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Db 741 GCTTTTAAAGTTTATCGCATTTGCAGCATTCGACCTTGGCCCTGACCAAAATATACCCAC 800
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 801 ATACCAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 860
Qy 221 ArgIleHisSerLeuGlyLysIleCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 861 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCTAGACACTTTAGAT 920
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 921 TTAATATTAATTAACCTTGATGAATCCCCACTGCATTTAGACACTCTCCAAACCTTAAA 980
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 981 GAACTAGGATTTTCATAGCAACAAATATCAGGTGATACCTGAGAAAGCATTTGTAGGCAAC 1040
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 1041 CCTTCTCTTATTACAAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 1100
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 1101 TTTTCAACATTTACCTGAATAGAACACTGACTGTGAATGTGCTGCACAAATAACTGAA 1160
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1161 TTTTCTGATTTAACTGGAACTGCAAAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1220
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1221 TCATCTCTTCTCAAAACCTGTCGAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1280
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1281 TACAACCTATTAGAAAGATTACCCAGTTTTTCAGTCTGCGCAAAAGCTTCAGAAATATGAC 1340
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1341 CTAAGACATATGAAATCTACGAAATTAAGTTGACATTTTCCAGCAGTTGCTTAGCCTC 1400

Qy 401 ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1401 CGATCGCTGAATTTGGCTTGGACAAAATTCCTATTATTACCCCAATGCATTTTCCACT 1460
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1461 TTGCCATCCCTAAATAAGCTGACCTATCGTCCAAACCTCCTGTCTGCTTTTCTTATAACT 1520
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1521 GGGTTACATGTTTAACTCACTTAAATAATTAACAGAAATCATGCTTTACAGAGCTTGATA 1580
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1581 TCATCTGAAACTTTTCAGAAACTCAGGTTATAGAAATGCCTTATGCTTACAGTCTGT 1640
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTyrAsnLysGlyAspAsn 500
Db 1641 GCATTTGGAGTGTGTGAGAAATGCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC 1700
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1701 AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1760
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1761 GACCTTGAAGATTTCTCGCTTGACTTTGAGAGAAAGCTGAAAGCCCTTCATTCAGTGGAG 1820
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIle 560
Db 1821 TGTTCACTTCCCCAGGCCCTTCAAACCTGTGAACACACTGCTTGATGGCTGGCTGATC 1880
Qy 561 ArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1881 AGAATGGAGTGTGGACCATAGCAGTTCTGCACATTACTTGTAAATGCTTTTGGTGACTTCA 1940
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1941 ACAGTTTTCAGATCCCTCTGTACATTTTCCCCCAATTAACGTTAATTTGGGGTCAATCGCA 2000
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2001 GCAGTGAACATGCTCACGGGAGTCTCCAGTCCCGTGTGCTGGCTGTGGATGCGTCACT 2060
Qy 621 PheGlySerPheAlaArgHisGlyAlaTyrTrpGluAsnGlyValGlyCysHisValIle 640
Db 2061 TTTGGCAGCTTTGACGACCATGGTGGTGGGAGAAATGGGGTGGTGGTGCATGTCAAT 2120
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 2121 GGTTTTTTGTCATTTTGTCTTCAAGATCATCTGTTCCTGCTTACTCTGGCAGCCCTG 2180
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2181 GAGCGTGGTGTCTCTGTGAAATAATTTGCAAAATTTGAAACGAAAGCTCCCATTTTCTAGC 2240
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2241 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGAGCATGGCCGCGAGTTCCCTCTG 2300
Qy 701 LeuGlyGlySerLysTyrGlyValaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2301 CTGGGTGGCAGCAAGTATGGCCCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGAGGCC 2360
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2361 AGCACCATTGGCTACATGGTGGCTCATCTGTGCTCAATTCCTTTGCTTCTCATGATG 2420
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr 760
Db 2421 ACCATTGCTCACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTCGAGATATTTGG 2480


```
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGAACAAAATTGCTATTATTACCCCAATGCAATTTCCACT 1260
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCCCTAAATAAGCTGGACCTATCGTCCAAACCTCTGCTGCTTTTCCCTATAACT 1320
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTAACTCACTTAAATTTACAGAAATCATGGCTTACAGAGCTTGATA 1380
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaIleLysGlnCysCys 480
Db 1381 TCATCTGAAATCTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTGT 1440
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1441 GCATTTGGAGTGTGTGAGAATGCTTATAGAAATGCTTAAATCAATGGAATAAAGGTGACAAC 1500
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCATGATGACGACCTTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1561 GACCTTGAAGATTTCTCTGCTTGACTTTGAGGAGACCTGGAAGCCCTTCATTCACTGAGTGCAG 1620
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisIleLeuLeuAspGlyTrpLeuIle 560
Db 1621 TGTTTCACTTCCCAGGCCCTTCAAAACCCCTGTGAACACCTGCTTGATGGCTGCTGCATC 1680
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1681 AGAATTTGGAGTGTGGACATAGCAGTTCCTGGCACTTACTTGTATGCTTTGGTGACTTCA 1740
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAATGTTAAATGGGGTCATCGCA 1800
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1801 GCAGTGAACATGCTTCACGGAGTCTCCAGTCCCGCTGCTGGCTGTGGATGGCTTCACT 1860
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTGCACACATGGTGGCTGGTGGGAGAAATGGGTTGGTTGCCATGTCAAT 1920
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1921 GGTTTTGTGTCATTTTGTCTTCAGATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAAATCAATTTTGGCTGTGGCCCTGTGGCCCTTGACCATGGCCGAGTTCCTG 2100
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2101 CTGGGTGGCAGCAAGTATGGGCCCTCCCTCTGCTGCTGCTTGGCTTTTGGGAGCCC 2160
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2161 AGCACCATGGCTACATGGTCTCATCTTGTCTCAATCCCTTTGCTTCTCATGATG 2220
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2221 ACCATTCCCTACACCAAGCTCTACTGCAATTTGACCAAGGAGACCTGGAGAAATATTGG 2280
Qy 761 AspCysSerMetVal-LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCy 780
```

```
Db 2281 GACTGCTCTATGAA-AAAAACACATTGCCCTGTGCTCTTCCACCAACTGCATCCTAAACTG 2339
Qy 780 sProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVa 800
Db 2340 CCTGTGGCTTCTCTGTGCTTCTCTCTTTAAATAAACCTTACATTTATCAGTCTGAAGT 2399
Qy 800 lIleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTy 820
Db 2400 AATTAAGTTTATCTTCTGTGTGTAGTCCACATCTCTGTCATGCTCAATCCCTTCTCTA 2459
Qy 820 rIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVa 840
Db 2460 CATCTGTTTCAATCCTCACCTTTAAGGAGGATCTGGTGAGCCTGAGAAAGCAAACTACGT 2519
Qy 840 lTrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysG1 860
Db 2520 CTGGACAAGATCAAAACACCAAGCTTGATGTCAATTAACCTGATGATGTGCAAAAACA 2579
Qy 860 nSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLe 880
Db 2580 GTCTGTGACTCAACTCAAGCCTTGGTAACCTTTACCAGCTCCAGCATCACTTATGACCT 2639
Qy 880 uProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSe 900
Db 2640 GCCTCCAGTTCGGTGGCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCCTC 2699
Qy 900 rValAlaPheValProCysLeu 907
Db 2700 TGTGGCATTTGTCCCATGTCTC 2721

RESULT 12
US-10-505-486-205
; Sequence 205, Application US/10505486
; Publication No. US20050118639A1
; GENERAL INFORMATION:
; APPLICANT: Takeda Chemical Industries, Ltd.
; TITLE OF INVENTION: Determination of a ligand
; FILE REFERENCE: P03-0006PCT
; CURRENT APPLICATION NUMBER: US/10/505,486
; CURRENT FILING DATE: 2004-08-20
; PRIOR APPLICATION NUMBER: JP 2002-45728
; PRIOR FILING DATE: 2002-02-22
; PRIOR APPLICATION NUMBER: JP 2002-213949
; PRIOR FILING DATE: 2002-07-23
; PRIOR APPLICATION NUMBER: JP 2002-298237
; PRIOR FILING DATE: 2002-10-11
; NUMBER OF SEQ ID NOS: 233
; SEQ ID NO 205
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Human
; US-10-505-486-205

Alignment Scores:
Pred. No.: 0 Length: 3438
Score: 806.00 Matches: 906
Percent Similarity: 99.78% Conservatives: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 88.86% Indels: 2
DB: 21 Gaps: 0

US-10-751-736-84 (1-907) x US-10-505-486-205 (1-3438)
Qy 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
Db 1 ATGACACACCTCCCGGCTCGGTGTGCTCTGCTTGCCTTGCCTGCTGCTGAGCTGGCGACC 60
Qy 21 GlyClySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Db 61 GGGGGCAGCTCTCCAGGCTGTGGTGTGCTGCTGAGGGGCTGCCCCACACACTGTCATGC 120
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
```


121	Db		GAGCCCGACGCGACGAGTGTGGCTCAGGGTGGACTGCTCCGACCTGGGGGCTCTCGAGCTG	180
61	Qy		ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln	80
181	Db		CTTTCGAACCTCAGCGTCTTCACTCTCTAGACCTCAGTAGAACACATCAGTCAG	240
81	Qy		LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn	100
241	Db		CTGCTCCGAATCCCTGCCAGTCTCCGCTTCTCGGAGGAGTTACGCTTCGCGGAAC	300
101	Qy		AlaLeuThrTyrIle-ProIysGlyAlaPheThrGlyLeuTyrSerLeuIysValLeuMe	120
301	Db		GCTCTGCATACGTTCCCAAGGAGGACATTCACTGGGCTTTACAGTCTTAAAGTTC	359
120	Qy		tLeuGlnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLe	140
360	Db		GCTGCAGATTAATCAGCTTAGACACGTTACCCACAGAAGCTCTGCAGAAATTCG	419
140	Qy		uGlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerG	160
420	Db		TCAATCCCTGGCTCTGGATGTAACCAACATCAGCTATGTGCCCAAGCTGTTTCA	479
160	Qy		YLeuHisSerLeuArgHisLeuTyrLeuAspAsnAlaLeuThrGluIleProValG	180
480	Db		CCTGCATTCCTGAGGACCTGTGGCTGGATGACAATCGCTTAACAGAAATCC	539
180	Qy		nAlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnIysIleHis	200
540	Db		GGCTTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTGCCCTGAACAAATAC	599
200	Qy		sIleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsn	220
600	Db		CATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGTAGTTCTACATCTCCATA	659
220	Qy		nArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAs	240
660	Db		TAGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTTAGAGA	719
240	Qy		pLeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeu	260
720	Db		TTTAAATTACAAATAACTTGATGAATTCCCACTGCAATTAGGACATCTTCCA	779
260	Qy		sGluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGly	280
780	Db		AGAACTAGGATTTTCATAGCAACAATATCAGTTCGATACCTGAGAAAGCATTT	839
280	Qy		nProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSer	300
840	Db		CCCTTCTCTATTATCAATATACATTTCTATGACAATCCCATCCATTTGTTGG	899
300	Qy		aPheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThr	320
900	Db		TTTTCAACATTTTACTGAACTTAAGAACACTGACTCTGAATGGTGCCCTCACAA	959
320	Qy		uPheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaG	340
960	Db		ATTTCTCTGATTTAACTGGAACCTGCAACCTGGAGAGCTGACTTTAACTGG	1019
340	Qy		eSerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeu	360
1020	Db		CTCATCTCTCTCCAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGTAG	1079
360	Qy		rTyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnIysLeuGlnIysIle	380
1080	Db		TTACAACCTATTAGAAGATTTTACCAGGTTTTTCAGTCTGCCAAAGCTTCAGA	1139
380	Qy		pLeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLe	400
1140	Db		CCTAAGACATTAATGAAATCTACGAAATTAAGTTGACACTTCCAGCAGTTGCT	1199
400	Qy		uArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSer	420

1200	CCGATCGCTGAAATTTGGCTTGGAAACAAAATTGCTATTATTATTCACCCCAATGCATTTTCCAC	1259
420	rLeuProSerLeuIleIysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleth	440
1260	TTTGCCATCCCTAATAAAGCTGCACCTATCGTCCAACCTCTGCTGCTCTTTTCCCTATAAC	1319
440	rGlyLeuHiIcGlyLeuThrHisLeuIysLeuThrGlyAsnHisAlaLeuGlnSerLeuI	460
1320	TGGGTACATGGTTTAACTCACTTAAATTTAAACAGGAATCATGCTCTTACAGAGCTTGAT	1379
460	eSerSerGluAsnPheProGluLeuIysValIleGluMetProTyAlaTyrGlnCysCy	480
1380	ATCATCTGAAAACTTTCCAGACTCAAGGTTATAGAAATGCCCTTATGCTTACCGATGCTG	1439
480	sAlaPheGlyValCysGluAsnAlaTyIysIleSerAsnGlnTrpAsnLysGlyAspAs	500
1440	TGCATTTGGAGTGTGTGAGAAATGCCATTAAGATTTCTTAATCAATGGAATAAAGGTGACAA	1499
500	nSerSerMetAspAspLeuHisIysIysAspAlaGlyMetPheGlnAlaGlnAspGluAr	520
1500	CAGCAGTATGCGACGACCTTCATAGAAGAGATGCTGGAATGTTTTCAGGCTCAAGATGAACG	1559
520	qAspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHiIcSerValG	540
1560	TGACCTTGAAGATTTCTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCA	1619
540	nCysSerProSerProGlyProPheLysProCysGluHiIcLeuLeuAspGlyTrpLeuI	560
1620	GTGTTACACTTCCCAGGCCCTTCAACCCCTGTGAACACCTGCTTGATGGCTGGCTGAT	1679
560	eArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSe	580
1680	CAGAAATGGAGTGTGGACCATAGCAGTTCTGGCAGCTTACTTGTAAATGCTTTGGTGACTTC	1739
580	rThrValPheArgSerProLeuTyIleSerProIleLysLeuLeuIcGlyValIleAl	600
1740	AACAGTTTTTCAGAAATCCCTCTGACATATTTCCCCCATTAACATGTTAATTTGGGGTCATGC	1799
600	aAlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheTh	620
1800	AGCAGTGAAACATGCTCAGCGGAGTCTCCAGTGGCGGTGCTGGCTGGTGGATGGATTCAC	1859
620	rPheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValI	640
1860	TTTTGGCAGCTTTGCGACACATCGTGCCTGCTGGGAGAAATGGGGTTGGTTGCCATGTCAT	1919
640	eGlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLe	660
1920	TGFTTTTTTGTCCATTTTTGCTTCAGAAATCATCTGTTTTCTCTGCTTACTCTGGCAGCCCT	1979
660	uGluArgGlyPheSerValLysTyIcSerAlaLysPheGluThrLysAlaProPheSerSe	680
1980	GGAGCGTGGGTTCTCTGTGAAATATTCGCAAAATTTGAAACGAAAGCTCCATTTTCTAG	2039
680	rLeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLe	700
2040	CCTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGGCTTGACCATGGCGCGAGTTCCCT	2099
700	uLeuGlyGlySerLysTyIcGlyAlaSerProLeuTyIcLeuProLeuProPheGlyGluPr	720
2100	GCTGGGTGGCAGCAGATATGGCGCCCTCCCTCTCTGCTGCCCTTGCCCTTTGGGGAGCC	2159
720	oSerThrMetGlyTyIcMetValAlaLeuIleLeuLeuAsnSerLeuTyIcPheLeuMetMe	740
2160	CAGCACCATGGGCTACATGTCCTCTCATCTTTGCTCAATTCCTTTTGCTTCTCTCATGAT	2219
740	tThrIleAlaTyIcThrLysLeuTyIcCysAsnLeuAspLysGlyAspLeuGluAsnIleTr	760
2220	GACCATTTGCCCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTCGAGAAATATTGG	2279
760	pAspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCyAlaIleAsnCy	780
2280	GGACTGCTATGGTAAAAACATATGCCCTTGTTGCTCTTCCACCACTGCATCTCAAACTG	2339


```
Qy 780 sProValAlaPheLeuSerPheSerLeuLeuAasnLeuThrPheIleSerProGluVa 800
Db 2340 CCCUGGCTTTCTTGCTCTCTCTCTTTAATAAAACCTTACATTATTCAGTCTGAGT 2399
Qy 800 lIleLysPheIleLeuLeuValValProLeuProAlaCysLeuAasnProLeuLeuTy 820
Db 2400 AATTAAGTTTATCTCTGTTGGTAGTCCACCTTCCTGCGATGTCCTCAATCCCTTCTCTA 2459
Qy 820 rIleLeuPheAasnProHisPheLysGluAaspLeuValSerLeuAargLysGlnThrTyVa 840
Db 2460 CATCTTGTTCAATCCTCACTTTAGAGGAGTCTGCTGAGCTCGAGAAAGCAAACTTACGT 2519
Qy 840 lTrpThrArgSerLysHisProSerLeuMetSerIleAasnSerAaspValGluLysG1 860
Db 2520 CTGACCAAGATCAAAACACCAAGCTTGATGTCNAATTAATCTGATGATGCGAAAACA 2579
Qy 860 nSerCysAaspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyRaspLe 880
Db 2580 GTCTGTGACTCAACTCAAGCTTGGTAACCTTTACCAGCTCCAGCATCACTTATGACCT 2639
Qy 880 uProProSerSerValProSerProAlaTyRProValThrGluSerCysHisLeuSerSe 900
Db 2640 GCCTCCAGTTCCGTGCGCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCCTC 2699
Qy 900 rValAlaPheValProCysLeu 907
Db 2700 TGTGGCATTGTCCCATGTCTC 2721

RESULT 13
US-10-295-027-483
; Sequence 483, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 483
```

```
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-483

Alignment Scores:
Pred. No.: 0 Length: 2651
Score: 621.00 Matches: 621
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.47% Indels: 0
DB: 17 Gaps: 0

US-10-751-736-84 (1-907) x US-10-295-027-483 (1-2651)
Qy 287 HisPheTyRAspAasnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu 306
Db 787 CATTTCTATGACAAATCCCAATTCCTGTTGGAGATCTGCTTTTCAACATTTACCTGAA 846
Qy 307 LeuArgThrLeuThrLeuAasnGlyAlaSerGlnIleThrGluPheProAaspLeuThrGly 326
Db 847 CTAAAGAACACTGACTCTGAATGGTGCCTCACAATAAATGAATTTCTGTATTTAACTGGA 906
Qy 327 ThrAlaAasnLeuGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThr 346
Db 907 ACTGCAAAACCTGGAGAGTCTGACTTTAACTGGAGACACAGATCTCATCTCTCTCTCAAACC 966
Qy 347 ValCysAasnGlnLeuProAasnLeuGlnValLeuAaspLeuSerTyRAsnLeuLeuGluAasp 366
Db 967 GTCTGCAATCAGTTACTTAATCTCCAGTGTAGATCTGTCTTACAACTTATTAAGAAT 1026
Qy 367 LeuProSerPheSerValCysGlnLysLeuGlnLysIleAaspLeuArgHisAasnGluIle 386
Db 1027 TTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTCAGACATAATGAATC 1086
Qy 387 TyrGluIleLysValAaspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAasnLeuAla 406
Db 1087 TAGCAAAATTAAGTTGACACTTTCAGACAGTTCTTAGCCTCCGATCGCTGGAATTTGGCT 1146
Qy 407 TrpAasnLysIleAlaIleHisProAasnAlaPheSerThrLeuProSerLeuLeuLys 426
Db 1147 TGGAAACAAAATGCTATTATTACCCCAATGCATTTTCCACTTTCATCCCTTAAATAAG 1206
Qy 427 LeuAaspLeuSerSerAasnLeuLeuSerPheProIleThrGlyLeuHisGlyLeuThr 446
Db 1207 CTGGACCTATCGTCCAACTCTCGTCTTTTCTTATATACTGGTTTACATGGTTTAACT 1266
Qy 447 HisLeuLysLeuThrGlyAasnHisAlaLeuGlnSerLeuIleSerSerGluAasnPhePro 466
Db 1267 CACTTAAATAAACAGAAATCATGCTTACAGAGCTTGATATCATCTGAAAACTTTCCA 1326
Qy 467 GluLeuLysValIleGluMetProTyRAlaTyRGlncysCysAlaPheGlyValCysGlu 486
Db 1327 GAACCTCAAGGTTATAGAAATGCTTATGCTTACAGTGTCTGTGCATTTGGAGTGTGAG 1386
Qy 487 AasnAlaTyRlysIleSerAasnGlnTrpAasnLysGlyAaspAasnSerSerMetAaspLeu 506
Db 1387 AATGCCCTATAAGATTCTAATCAATCGAATAAAGGTGACAAACAGCAGTATGGACGCTT 1446
Qy 507 HisLysLysAaspAlaGlyMetPheGlnAlaGlnAaspGluArgAaspLeuGluAaspPheLeu 526
Db 1447 CATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTCTG 1506
Qy 527 LeuAaspPheGluGluAaspLeuLysAlaLeuHisSerValGlnCysSerProSerProGly 546
Db 1507 CTTGACTTTTGAGGAAGACCTGGAAGCCCTTCATTTCAGTGGAGTGTTCACCTTCCCAAGGC 1566
Qy 547 ProPheLysProCysGluHisLeuLeuAaspGlyTrpLeuIleArgIleGlyValTrpThr 566
Db 1567 CCCTTCAAAACCTGTGAACACCTGTTGATGGCTGGCTGATCAGAAATTGGAGTGTGACC 1626
Qy 567 IleAlaValLeuAlaLeuThrCysAasnAlaLeuValThrSerThrValPheAsgSerPro 586
```


Db	1027	TTACCCAGTTTTTCAGTCTGCCAAGAGCTTCAGAAAATTTGACCTAAGACATATGAATC	1086
Qy	387	TyrGluLeuLysValAspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAAsnLeuAla	406
Db	1087	TACGAAATTTAAAGTTGACACTTTTCAGCAGTTGCTTAGCCCTCCGATCGCTGAATTTGGCT	1146
Qy	407	TrpAsnLysIleAlaIleHleHisProAsnAlaPheSerThrLeuProSerLeuLleLys	426
Db	1147	TGGAACAAAATTTGCTATTATTCACCCCAATTCATTTTCCACTTTGCCATCCCTAATAAAG	1206
Qy	427	LeuAspLeuSerSerAsnLeuLeuSerPheProIleThrGlyLeuHisGlyLeuThr	446
Db	1207	CTGGACCTATCGTCCAACTCTCTGCTCTTTTCTATTAAGTGGTTACATGGTTTAACT	1266
Qy	447	HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuLeuSerSerGluAAsnPhePro	466
Db	1367	CACITTAATAATTAACAGGAATCATGCCCTTACAGAGCTTGATATCATCTGAATACTTTCCA	1326
Qy	467	GluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGlu	486
Db	1327	GAACCTAAGGTTATAGAAATGCCCTTATGCTTTACCCAGTGTGTGCATTTGGAGTGTGAG	1386
Qy	487	AsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu	506
Db	1387	AATGCCCTATAAGATTTCTAATCAATGGAATTAAGGTGACCAACAGCAGTATGGACGACTT	1446
Qy	507	HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAAspPheLeu	526
Db	1447	CATAAGAAAGATGCTGGAATGTTTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGT	1506
Qy	527	LeuAspPheGluGluAspLeuLysAlaIleuHisSerValGlnCysSerProSerProGly	546
Db	1507	CTTGACCTTTGAGGAAGACCTTGAAAGCCCTTCATTTCAGTCAGTGTTCACCTTCCCCAGGC	1566
Qy	547	ProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThr	566
Db	1567	CCCTTCAAAACCTGTGAACACCTGCTTGAATGGCTGGCTGATCAGAAATGGAGTGGACC	1626
Qy	567	IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro	586
Db	1627	ATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCAACAGATTTTCAGATCCCT	1686
Qy	587	LeuTyrIleSerProLleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr	606
Db	1687	CTGFACATTTCCCCCAATTAACCTGTTAATTTGGGTGATCGCAGCAGTGAACATGCTCACG	1746
Qy	607	GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg	626
Db	1747	GGAGTCTCCAGTGCCTGGCTGGCTGGTGGATGCGTTCACTTTTGGCAGCTTTGGCAGCA	1806
Qy	627	HisGlyAlaTrpTrpGluAAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePhe	646
Db	1807	CATGCTGCTGGTGGGAGAAATGGGTTGGTTGCCATGTCAATTTGTTTTTGTCCATTTT	1866
Qy	647	AlaSerGluSerSerValPheLeuLeuThrIleuAlaLeuGluArgGlyPheSerVal	666
Db	1867	GCTTCAGAAATCATCTGTTTTCTCTACTCTGCAGCCCTGGAGCGTGGGTTCTCTGTG	1926
Qy	667	LysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValIleIleLeu	686
Db	1927	AAATATTCGAAAATTTGAACAGAAAGCTCCATTTTCTAGCCCTGAAAGTAATCATTTTG	1986
Qy	687	LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyr	706
Db	1987	CTCTGTGCCCTCTCTGCTGCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2046
Qy	707	GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet	726
Db	2047	GGCCGCCCTCTCTGCTGCCCTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2106
Qy	727	ValAlaLeuLeuLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLys	746

Db	2107	GTGCTCTCATCTCTGCTCAATTCCTTTGCTTCTCATGATGACCAATGGCTACACCAAG	2166
Qy	747	LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLys	766
Db	2167	CTCTACTGCAATTTGGACACAGGAGACCTTGGAGAAATATTTGGGACTGCTCTATGGTAAAA	2226
Qy	767	HisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSer	786
Db	2227	CACATTTGCCCTGTTGCTCTTTCACCAACTGCATCTTAACTGCCCTGTGGTTCCTTGTCC	2286
Qy	787	PheSerSerLeuLeuLeuThrPheIleSerProGluValIleLysPheIleLeuLeu	806
Db	2287	TTCTCTCTTTTAATAAACCTTACATTTATCAGTCTCTGAAGTAATTAAGTTTATCTTCTG	2346
Qy	807	ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis	826
Db	2347	GTGGTAGTCCCACTTCTCTGCATGCTCAATCCCCTTCTCTACATCTTGTTCATTCCTCAC	2406
Qy	827	PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTrpThrArgSerLysHis	846
Db	2407	TTTAAGAGAGATCTGGTGACCTTGAGAAACCAACCTACGCTTGGACACAGATCAAAACAC	2466
Qy	847	ProSerLeuMetSerIleAsnSerAspValGluLysGlnSerCysAspSerThrGln	866
Db	2467	CCAAGCTTTGATGTCAATTAACCTCTGATGATGTGCAAAAACAGTCTCTGTGACTCAACTCAA	2526
Qy	867	AlaLeuValThrPheThrSerSerIleThrTyrAspLeuProProSerSerValPro	886
Db	2527	GCCTTGGTAACTTTTACCAGCTCCAGCATCACATATGACCTGCCCTCCAGTTCGGTGCCA	2586
Qy	887	SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys	906
Db	2587	TCACCACTTATCCAGTACTGAGAGCTGCCATCTTCTCTGTGGCATTTGTCCCATGT	2646
Qy	907	Leu 907	
Db	2647	CTT 2649	
RESULT 15			
US-10-173-999-27			
; Sequence 27, Application US/10173999			
; Publication No. US20040005563A1			
; GENERAL INFORMATION:			
; APPLICANT: Mack, David H.			
; APPLICANT: Gish, Kurt C.			
; APPLICANT: Eos Biotechnology, Inc.			
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions			
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian			
; TITLE OF INVENTION: Cancer			
; FILE REFERENCE: 018501-002420US			
; CURRENT APPLICATION NUMBER: US/10/173,999			
; CURRENT FILING DATE: 2002-06-17			
; PRIOR APPLICATION NUMBER: US 60/299,234			
; PRIOR FILING DATE: 2001-06-18			
; PRIOR APPLICATION NUMBER: US 60/315,287			
; PRIOR FILING DATE: 2001-08-27			
; PRIOR APPLICATION NUMBER: US 60/350,666			
; PRIOR FILING DATE: 2001-11-13			
; PRIOR APPLICATION NUMBER: US 60/372,246			
; PRIOR FILING DATE: 2001-04-12			
; NUMBER OF SEQ ID NOS: 163			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 27			
; LENGTH: 2651			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
US-10-173-999-27			
Alignment Scores:			
Pred. No.: 0			
Score: 621.00			
Percent Similarity: 100.00%			
Best Local Similarity: 100.00%			
Length: 2651			
Matches: 621			
Conservative: 0			
Mismatches: 0			

Query Match:	68.47%	Indels:	0
DB:	17	Gaps:	0
US-10-751-736-84 (1-907) x US-10-173-999-27 (1-2651)			
QY	287	HisPheThrSerProLeuGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu	306
DB	787	CATTTCTATGACAAATCCATCCAAATTTGTTGGGAGATCTGCTTTTCAACATTTACCTGAA	846
QY	307	LeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGly	326
DB	847	CTAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAATTTCTTGATTTAACTGGA	906
QY	327	ThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThr	346
DB	907	ACTGCAAAACCTGGAGAGTCTGACTTTAACTGGAGCAGATCTCATCTCTCTCTCAAAAC	966
QY	347	ValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSerThrAsnLeuLeuGluAsp	366
DB	967	GTCTGCAATCAGTTACCTAATCTCCAAAGTGTAGATCTGTCTTACAACTATTAGAAAT	1026
QY	367	LeuProSerPheSerValCysGlnLeuGlnLysIleAspLeuArgHisAsnGluIle	386
DB	1027	TTACCCAGTTTTTCAGTCTGCCAAAGCTTCAGAAATTCACCTTAGACATATATGAATC	1086
QY	387	TyrGluIleLysValAspThrPheGlnLeuLeuSerLeuLeuArgSerLeuAsnLeuAla	406
DB	1087	TACGAAATTAAGTTGACATTTCCAGCAGTTGCTTAGCCCTCCGATCGCTGAATTTGGCT	1146
QY	407	TrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThrLeuProSerLeuIleLys	426
DB	1147	TGGAACAAATTCGTAATTTACCCCAATTCGCAATTTTCCACTTTGCCATCCCTTAATAAG	1206
QY	427	LeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThrGlyLeuHisGlyLeuThr	446
DB	1207	CTGGACCTATCGTCCAACCTCTCTGCTCTTTTCTTATATACTGGGTACATGGTTAACT	1266
QY	447	HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPhePro	466
DB	1267	CACCTAAATTAACAGAAATCATGCTTACAGAGCTTGATATCATCTGAAATCTTTCCA	1326
QY	467	GluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGlu	486
DB	1327	GAATCAAGGTATAGAAATGCTCTATGCTTACAGTGTCTGCTATTTGGAGTGTGAG	1386
QY	487	AsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu	506
DB	1387	AATGCCATTAAGATTTCTAATCAATGGAAATTAAGGTGACCAACAGCAGTATGGACCTT	1446
QY	507	HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeu	526
DB	1447	CATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAACGTGACCTTGAAGATTTCTG	1506
QY	527	LeuAspPheGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGly	546
DB	1507	CTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGAGTGTTCACCTTCCCCAGGC	1566
QY	547	ProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThr	566
DB	1567	CCCTTCAAAACCTGTGAACACCTGCTTGATGGCTGGCTGATCAGAAATGGAGTGTGACC	1626
QY	567	IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro	586
DB	1627	ATAGCAGTTCTGCGACTTACTTGTAAATGCTTTGTGTGACTTCAACAGTTTTCAGATCCCT	1686
QY	587	LeuTyrIleSerProLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr	606
DB	1687	CTGTACATTTTCCCCCATTAATCTTTAAATTTGGGTCTATCGCAGCAGTGAACATGCTCACG	1746
QY	607	GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg	626
DB	1747	GGAGTCTCCAGTCCGCTGGCTGGTGGTGGATCGCTTCACTTTTGGCAGCTTTTGCACGA	1806

Search completed: July 12, 2005, 13:53:20
Job time : 3118 secs

QY	627	HisGlyAlaIleTrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePhe	646
DB	1807	CATGGTGCCTGGTGGGAGAAATGGGGTTGGTGCATGTCAATGGTTTTTTTGCATTTTT	1866
QY	647	AlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerVal	666
DB	1867	GCTTCAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTCTGTG	1926
QY	667	LysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValIleLeuLeu	686
DB	1927	AAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTGAAAGTAATCATTTG	1986
QY	687	LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyr	706
DB	1987	CTCTGTGCCCTGTGCTGCTTGACCATGGCGCGAGTTCCCTGTCTGGGTGGCAGCAAGTAT	2046
QY	707	GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet	726
DB	2047	GGCGCCTCCCTCTCTGCTGCTTTCCTTTTGGGAGCCAGCACCATTGGGCTACATG	2106
QY	727	ValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLys	746
DB	2107	GTGCTCTCATCTTGTCAATTCCTTGTCTCTCATGATGACCATTTGCTACACCAAG	2166
QY	747	LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLys	766
DB	2167	CTCTACTCAATTTGGACAAGGAGACCTGGAGAATATTTGGGAGCTGCTCTATGGTAAAA	2226
QY	767	HisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSer	786
DB	2227	CATTTGCCCTGTGCTCTTCCAACTGCATCTCAATCCCTTCTTACATCTTGTGGCTTTCTG	2286
QY	787	PheSerSerLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeu	806
DB	2287	TTCTCTCTTTTAATAAACCTTACATTTATCAGTCTCTGAAGTAATTAAGTTTATCTCTG	2346
QY	807	ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis	826
DB	2347	GTGGTAGTCCCACTTCTGTCATGCTCAATCCCTTCTCTACATCTTGTTCATCTCCAC	2406
QY	827	PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTrpThrArgSerLysHis	846
DB	2407	TTTAAGAGGATCTGTGAGCCTGAGAAAGCAAACTACGCTCTGGACNAGATCAAAACAC	2466
QY	847	ProSerLeuMetSerIleAsnSerAspValGluLysGlnSerCysAspSerThrGln	866
DB	2467	CCAAGCTTGATGTCAATTAACCTCTGATGATGTCGAAAAACAGTCTCTGACTCAACTCAA	2526
QY	867	AlaLeuValThrPheThrSerSerIleThrTyrAspLeuProProSerSerValPro	886
DB	2527	GCCTTGGTAACCTTTTACCAGCTCCAGCATCACATTATGACCTGCCCTCCAGTTCGGTCCA	2586
QY	887	SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys	906
DB	2587	TCACCAGCTTATCCAGTGACTGAGAGTGCATCTTCTCTGTGGCATTTTGTCCCATGT	2646
QY	907	Leu 907	
DB	2647	CTT 2649	

This Page Blank (uspto)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2005, 07:53:52 ; Search time 6975 Seconds
(without alignments)
4949.719 Million cell updates/sec

Title: US-10-751-736-84

Perfect score: 907

Sequence: 1 MDSRLGVLLSLPVLQLAT.....PAYPTVTSCHLSVAVFVCL 907

Scoring table:

OLIGO	Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 34239544 seqs, 19032134700 residues

Word size: 1

Total number of hits satisfying chosen parameters: 68477535

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool_US10751736/runat_12072005_083641_22183/app_query.fasta_1.1095
-DB=EST -QFMT=fastap -SUFFIX=oli.rst -MINMATCH=0 -1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751736 @CGN 1 1 5210 @runat_12072005_083641_22183 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST:*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_hc:*
4:	gb_est3:*
5:	gb_est4:*
6:	gb_est5:*
7:	gb_est6:*
8:	gb_gss1:*
9:	gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	764	84.2	2724	9 AY411732	Homo sapi
2	316	34.8	1060	1 AL541959	AL541959
3	235	25.9	800	4 BG196708	BG196708
4	179	19.7	590	6 CD617354	CD617354
5	176	19.4	873	5 BX403219	BX403219
6	158	17.4	475	5 BX107244	BX107244
7	157	17.3	1104	1 AL530798	AL530798
8	156	17.2	472	1 AA460529	AA460529
9	156	17.2	556	5 BQ323949	BQ323949

10	149	15.4	671	4	BM795259
11	137	15.1	634	7	CF177761
12	122	13.5	374	1	AI367723
13	118	13.0	358	1	AI699236
14	111	12.2	539	1	AI697103
15	98	10.8	2593	9	AY411733
16	95	10.5	494	1	AA424098
17	91	10.0	576	5	BX499910
18	83	9.2	276	6	CD617346
19	83	9.2	276	6	CD617347
20	83	9.2	276	6	CD617348
21	83	9.2	276	6	CD617349
22	83	9.2	276	6	CD617353
23	81	8.9	268	6	CD617342
24	81	8.9	271	6	CD617352
25	78	8.6	250	6	CD617356
26	74	8.2	222	6	CD617357
27	74	8.2	222	6	CD617358
28	72	7.9	273	6	CD617343
29	65	7.2	940	7	CK449409
30	64	7.1	330	5	BX357656
31	58	6.4	224	6	CD617359
32	57	6.3	432	8	AQ076504
33	56	6.2	275	6	CD617344
34	51	5.6	627	2	BB613285
35	51	5.6	763	7	CF742674
36	51	5.6	2724	9	AY411734
37	51	5.6	4796	3	AK047873
38	48	5.3	492	6	CD617355
39	46	5.1	397	1	AI539105
40	45	5.0	275	6	CD617350
41	45	5.0	275	6	CD617351
42	42	4.6	698	5	BF143829
43	41	4.5	542	1	AI962439
44	39	4.3	280	6	CD617345
45	39	4.3	396	1	AL119316

ALIGNMENTS

RESULT 1	AY411732	2724 bp	DNA	linear	GSS 16-DEC-2003
LOCUS	AY411732	Homo sapiens GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.			
DEFINITION	AY411732				
ACCESSION	AY411732.1	GI:39767700			
VERSION	GSS.				
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2724)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 2724)				
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	location/Qualifiers				
source	1. .2724				


```

RESULT 3
BG196708
LOCUS RST15934 Athersys RAGE Library Homo sapiens cDNA, mRNA linear EST 21-APR-2001
ACCESSION BG196708
VERSION BG196708.1 GI:13718395
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 800)
AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
Cain,S., Leventhal,C., Thornton,M., Ramachandran,R.,
Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Boozer,S.,
Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
Offenbacher,J., Danzig,J. and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random
activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 481.
FEATURES
source
1..800
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."
ORIGIN
Alignment Scores:
Pred. No.: 4,57e-233 Length: 800
Score: 235.00 Matches: 235
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.91% Indels: 0
DB: 4 Gaps: 0
US-10-751-736-84 (1-907) x BG196708 (1-800)
QY 404 AnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThrLeuProSer 423
Db 1 AATTTGGCTTGAACAAAATTGCTATTATACCCCAATGCATTTCCACTTTGGCATCC 60
QY 424 LeuIleLysLeuAspLysSerSerAsnLeuLeuSerPheProIleThrGlyLeuHis 443
Db 61 CTAATAAGCTGGACCTATCGTCCAACTCCTGCTCTTTTCTTAATACCTGGTTACAT 120
QY 444 GlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerIleSerSerGlu 463
Db 121 GGTTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATATCATCTGAA 180
QY 464 AsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGly 483
Db 181 AACTTTCCAGAACTCAAGGTTATAGAAATGCCTTTATGCTTACAGTGTGCATTGGA 240
QY 484 ValCysGluAsnAlaTyrLysIleSerAsnGlnTTPAsnLysGlyAspAsnSerSerMet 503
Db 241 GTGTGTGAGAATGCCTATAAGATTCTTAATCAATGAATAAGGTGACAAACAGCATG 300
QY 504 AspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGlu 523
Db 301 GACGACCTTCATAGAAAAGATGCTGGAATGTTTACGGCTCAAGATGAACGTCACCTTGA 360
QY 524 AspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGlnCysSerPro 543
Db 361 GATTTCCCTGCTTGACTTTGAGGAGACCTGAAAGCCCTTCATTCACTGACAGTTCACCT 420
QY 544 SerProGlyProPheLysProCysGluHisLysLeuLeuAspGlyTTPLeuIleArlGly 563
Db 421 TCCCCAGGCCCTTCAAAACCTGTGAACACCTGCTGATGCTGGCTGATCAGAAATTGA 480
QY 564 ValTTPThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPhe 583
Db 481 GTGTGGACCATAGCAGTCTGCGACTTACTTGTAAATGCTTTGGTGACTTCAACAGTTTC 540
QY 584 ArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAlaValAsn 603
Db 541 AGATCCCTCTGTACATTTCCCCCATTAACCTGTAATTGGGTCATCGCAGCAGTGAAC 600
QY 604 MetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySer 623
Db 601 ATGCTCACGGAGTCTCCAGTCCGCTGCTGGTGGTGGATGCCGTTCACTTTGGCAGC 660
QY 624 PheAlaArgHisGlyAlaTTPTrpGluAsnGlyValGlyCysHis 638
Db 661 TTTGCACGACATGGTGGCTGCTGGTGGAGAAATGGGGTTGGTTGCCAT 705
RESULT 4
CD617354 590 bp mRNA linear EST 12-JAN-2004
LOCUS 55041415H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD617354
VERSION CD617354.1 GI:40265619
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 590)
AUTHORS Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.I.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Fu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
Location/Qualifiers
source
1..590
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"
ORIGIN
Alignment Scores:
Pred. No.: 6,61e-175 Length: 590
Score: 179.00 Matches: 179
Percent Similarity: 100.00% Conservatve: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 19.74% Indels: 0
DB: 6 Gaps: 0
US-10-751-736-84 (1-907) x CD617354 (1-590)
QY 98 AlaGlyAsnAlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLys 117
Db 16 GCGGAAACGCTCTGACATACATTCACCAAGGAGCATTCCTGGCGCTTTACAGCTTAA 75

```


QY 118 ValLeuMetLeuGlnAnAnGlnLeuArgHisValProThrGluAlaLeuGlnAnLeu 137
 Db 76 GTTCTTATGCTGCAGATAATCAGCTAAGACACGTACCCACAGAAGCTCTGCAGATTG 135
 QY 138 ArgSerLeuGlnSerLeuArgLeuAspAlaAnHisIleSerTyrrValProSerCys 157
 Db 136 CGAAGCTTCATCCCTGGCTGGATGTAAACCATCAGCTATGTGTGCCCCAGCTGT 195
 QY 158 PheSerGlyLeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIle 177
 Db 196 TTTCAAGTGGCTGCATTCCTGAGGACCTGTGGTGCATGACATCGTTAAACAGAAATC 255
 QY 178 ProValGlnAlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAnIys 197
 Db 256 CCGCTCCAGGCTTTTAAAGATTTATCGGATTCGCAAGCCATGACCTTGGCCCTGAAACAA 315
 QY 198 IleHisIleProAspTyrrAlaPheGlyAnLeuSerSerLeuValValLeuHisLeu 217
 Db 316 ATACACCATACCATGAGCTATGCTTTGGAAACCTTCAGCTTGGTAGTTCTACATCTC 375
 QY 218 HisAnAsnArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGlu 237
 Db 376 CATAACAATAGAAATCCCTCGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAG 435
 QY 238 ThrLeuAspLeuAnTyrrAnAnLeuAspGluPheProThrAlaIleArgThrLeuSer 257
 Db 436 ACTTTAGATTTAAATTTACAAATACCTTTGATGAATTCCTCCACTGCAATTAGGACACTCTCC 495
 QY 258 AsnLeuLysGluLeuGlyPheHisSerAnAnIleArgSerIleProGluLysAla 276
 Db 496 AACCTTAAAGAACTAGGATTTTATAGCAACAATATCAGTGCATACCTGAGAAGCA 552

RESULT 5
 BX403219
 LOCUS
 DEFINITION
 BX403219 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 Clone CS0D1029YG09 5-PRIME, mRNA sequence.
 BX403219
 BX403219.2 GI:46875014
 EST.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1. (bases 1 to 873)
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30619059.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 10101.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?8=CS2AAW3ZE04_AW92A03_1&c=10101.f
 Location/Qualifiers
 1. .873
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1029YG09"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 1,3e-171 Length: 873
 Score: 176.00 Matches: 176
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 19.40% Indels: 0
 DB: 5 Gaps: 0

US-10-751-736-84 (1-907) x BX403219 (1-873)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
 Db 138 ATGGACACCTCCCGGCTCGTGTGCTCTGCTCTTTCCTGTGTGCTGTGAGTGGCGACC 197
 QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
 Db 198 GGGGGGAGCTCTCCAGGCTGTGTGCTGAGGGGCTGCCACACACTGTCTATTGC 257
 QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
 Db 258 GAGCCCGACGGCAGGATGTTGCTCAGGGGTGGACTGCTCCGACCTGGGGCTCTCGAGCTG 317
 QY 61 ProSerAnLeuSerValPheThrSerTyrrLeuAspLeuSerMetAnAnIleSerGln 80
 Db 318 CTTTCCAACTCAGGCTCTTCACTCCTAGACCTCAGTATGAACACATCAGTCAG 377
 QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
 Db 378 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCGAGGAGTTACGTTTTCGGGAAAC 437
 QY 101 AlaLeuThrTyrrIleProLysGlyAlaPheThrGlyLeuTyrrSerLeuLysValLeuMet 120
 Db 438 GCTCTGACATACATTTCCCAAGGAGGACATTCACCTGGCCTTTACAGTCTTAAAGTCTTATG 497
 QY 121 LeuGlnAnAnGlnLeuArgHisValProThrGluAlaLeuGlnAnLeuArgSerLeu 140
 Db 498 CTGCAGAAATATCAGCTAAGACACAGTACCCACAGAAGCTCTGCAGAAATTTGGAGACCTT 557
 QY 141 GlnSerLeuArgLeuAspAlaAnHisIleSerTyrrValProProSerCysPheSerGly 160
 Db 558 CAATCCCTGCGTCTGGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTTCAGTGGC 617
 QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGlu 176
 Db 618 CTGCATTCCTGAGGACCTGTGGCTGGATGACATGCGTTAACAGAA 665

RESULT 6

BX107244

LOCUS

DEFINITION

BX107244 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone

IMAGp998N171964; IMAGE:796524, mRNA sequence.

ACCESSION

BX107244

VERSION

BX107244.1 GI:27847308

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 475)

Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,

Radelof, U., Schneider, D. and Korn, B.

Human Unigeneset - RZPD3

Unpublished (2003)

Contact: Ina Rolfs

COMMENT

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAGp998N171964.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

Human Unigeneset - RZPD3 (RZPDLIB No.972)

http://www.rzpd.de/CloneCards/cgi-

bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs

BX107244 475 bp mRNA linear EST 06-FEB-2003
 BX107244 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
 IMAGp998N171964; IMAGE:796524, mRNA sequence.

ACCESSION

BX107244

VERSION

BX107244.1 GI:27847308

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 475)

Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,

Radelof, U., Schneider, D. and Korn, B.

Human Unigeneset - RZPD3

Unpublished (2003)

Contact: Ina Rolfs

COMMENT

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH

Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD; IMAGp998N171964.

RZPDLIB; I.M.A.G.E. cDNA Clone Collection;

Human Unigeneset - RZPD3 (RZPDLIB No.972)

http://www.rzpd.de/CloneCards/cgi-

bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel.: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de

This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCCACAGGAACACGATGAC.

FEATURES

source
Location/Qualifiers

1..475
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:988N171964 ; IMAGE:796624"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus Nb2HF8 9w"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5'
TGTTACCAATCGAAGCGGAGCGCCGCTTAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:
Pred. No.: 3,85e-153 Length: 475
Score: 158.00 Matches: 158
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 17.42% Indels: 0
DB: 5 Gaps: 0

US-10-751-736-84 (1-907) x BX107244 (1-475)

QY 566 ThrLeuAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSer 585
DB 2 ACCATGACAGTTCTGGCATTACTGTAAATGCTTTGGTGACTTCAACAGTTTTCAGATCC 61
QY 586 ProLeuTyrIleSerProIleLeuLeuLeuGlyValIleAlaAlaValAsnMetLeu 605
DB 62 CCTCTGTACATTCCCCCATTAATCTGTTAATTTGGGGTCATCGCAGCAGTGAACATGCTC 121
QY 606 ThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAla 625
DB 122 ACGGAGTCTCCAGTCCGCTGCTGGTGTGGATGGTTCACCTTTTGGCAGCTTTCGA 181
QY 626 ArgHisGlyAlaTrrPrluAsnGlyValGlyCysHisValIleGlyPheLeuSerIle 645
DB 182 CGACATGGTCCCTGGTGGGAGAAATGGGGTTGGTGGCCATGTCTATGGTTTGTGCCATT 241
QY 646 PheAlaSerGluSerSerValPheLeuLeuThrIleAlaAlaLeuGluArgGlyPheSer 665
DB 242 TTTGCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTCCT 301
QY 666 ValIysTyrSerAlaLysPheGluThrIlysAlaProPheSerSerLeuLysValIleIle 685
DB 302 GTGAATATCTGCAAAATTTGAACAGAAAGCTCCATTTCTAGCCTGAAGTATCATTT 361
QY 686 LeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLys 705
DB 362 TTGCTCTGTGCCCTCTGCGCTTGAACCATGGCGCAGTTCCTCTGCTGGTGGCAGCAAG 421
QY 706 TyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluProSerThrMet 723
DB 422 TATGGCGCCCTCCCTCTCTGCGCTTTCCTTTTGGGGAGCCGCCACCAATG 475

RESULT 7
AL530798

LOCUS AL530798 1104 bp mRNA linear EST 24-MAR-2004
DEFINITION AL530798 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
cDNA clone CS0DD008YF22 5-PRIME, mRNA sequence.
ACCESSION AL530798
VERSION AL530798.3 GI:45705748
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 1104)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:31068631.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Crémieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7021.f

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?s=CS0DD008DC11QPI&c=7021.f.
Location/Qualifiers
1..1104
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD008YF22"
/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES

source

ORIGIN

Alignment Scores:
Pred. No.: 9,85e-152 Length: 1104
Score: 157.00 Matches: 312
Percent Similarity: 98.42% Conservative: 0
Best Local Similarity: 98.42% Mismatches: 2
Query Match: 17.31% Indels: 5
DB: 1 Gaps: 0

US-10-751-736-84 (1-907) x AL530798 (1-1104)

QY 334 ThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThrValCysAsnGlnLeuProAsn 353
DB 2 ACTTTAACTGGAGCAGACATCTCATCTCTCTCAACCGCTGCAATCAGTTACTTAAT 61
QY 354 LeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAspLeuProSerPheSerValCys 373
DB 62 CTCCAAGTGCTAGATCTGCTTACAACTATTAGAAATTTACCCAGTTTTCAGTCTGC 121
QY 374 GlnLysLeuGlnLysIleAspLeuArgHisAsnGluIleTyrGluIleLysValAspThr 393
DB 122 CAAAAGCTTCAGAAAAATTGACCTTAAGACATAATGAATCTACGAAATTAAGTTGACACT 181
QY 394 PheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAlaTrrPsnIleAlaIleIle 413
DB 182 TTCCAGCAGTTGCTTAGCTCCGATCGCTGAATTTGGCTTGGAAACAAATATGCTATTATT 241
QY 414 HisProAsnAlaPheSerThrLeuProSerLeuIleLysLeuAspLeuSerSerAsnLeu 433
DB 242 CACCCCAATGCATTTTCCACTTTGGCATCCCTAATAAAGCTGGACCTATCGTCCAACTC 301
QY 434 LeuSerSerPheProIleThrGlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsn 453

Db 302 CTGCTGCTCTTCTTATTAAGTGGTTACATGGTTTAACTCACTTAAATTAACAGGAAT 361
 QY 454 HsAlaLeuGlnSerLeuIleSerSerGluAenPheProGluLeuIleValleGluMet 473
 Db 362 CATGCTTACAGAGCTTGATATCATCTGAAGAACTTCCAGAACTCAAGGTTATAGAAATG 421
 QY 474 Pro-TyralaTyrglnCysCysAlaPheGlyValCysGlu-AenAlaTyrllys-IleSer 492
 Db 422 CTTATATGCTTACAGAGTGGTGTGCTATTTGAGTGTGTTA-GAATGCTTATWA-GATTTCT 479
 QY 493 AenGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeuHsIleLysLysAspAlaGly 512
 Db 480 NATCAATGGAATAAAGTGCACACAGCAGTATGGAGCAGCTTCATAGAAGAAGATGCTGA 539
 QY 513 MetPheGlnAlaGlnAspGluArgAspLeuGluAaspPheLeuLeuAaspPheGluLeuAep 532
 Db 540 ATGTTTTCAGGCTCAAGATGAAGCTGACCTTGAAGATTTCTGCTGACTTTGAGGAAGAC 599
 QY 533 LeuLysAlaLeuHsSerValGlnCysSerProSerProGlyProPheLysProCysGlu 552
 Db 600 CTGAAGCCCTTCATTCAGTCAGTGTGTTCACTTCCCGAGGCCCTTCAACCCCTGTGA 659
 QY 553 HsLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpTrpIleAlaValLeuAlaLeu 572
 Db 660 CACCTGCTTGATGGCTGGCTGATCAGATTTGAGTGTGACCATAGCAGTCTTGGCACTT 719
 QY 573 ThrCysAsnAlaLeuValThrSerThrValPheAerGserProLeuTyrlleSerProle 592
 Db 720 ACTTGTAATGCTTTGCTGACTTCAACAGTGTTCAGATCCCTCTGTATACATTTCCCCCAT 779
 QY 593 LysLeuLeuIleGlyValleAlaAlaValAenMetLeuThrGlyValSerSerAlaVal 612
 Db 780 AAACGTGTTAAATTTGGGGTTCATCCAGCAGTGAACATGCTCACGGAGTTCACAGTCCCGTG 839
 QY 613 LeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArgHsIleGlyValaTrpGlu 632
 Db 840 CTGGCTGGTGTGATGGCTGCTTCACTTTGGCAGCTTTGACAGCATGTTGGCTGGTGGGAG 899
 QY 633 AenGlyValGlyCysHsValleGlyPheLeuSerIlePheAla 647
 Db 900 AATGGGTTGGTTCATGTCATGTTGGTTTGGTTCATTTTGGC 944

RESULT 8

AA460529

LOCUS

DEFINITION xx68c09.r1 Soares total fetus Nb2HF8 9w Homo sapiens cDNA clone
 IMAGE:796624 5' similar to WP:CS0H2.1 CE05479 STEROID HORMONE

RECEPTOR ; mRNA sequence.

ACCESSION

AA460529

VERSION

AA460529.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
 Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
 Moore, B., Schellenberg, K., Steptoe, M., Tan, P., Theising, B.,
 White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE

JOURNAL

COMMENT

Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

This clone is available royalty-free through LILNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28ml3 rev2 ET from Amersham.

FEATURES

source

1. .472

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="GDB:6040989"

/db_xref="taxon:9606"

/clone="IMAGE:796624"

/dev_stage="8-9 weeks"

/lab_host="DH10B"

/clone_lib="Soares total_fetus Nb2HF8 9w"

/note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from pooled 8-9 week
 (total) fetus material with a Not I - oligo(GT) primer [5'
 TGTTACCAATCTGAGTGGAGCGCCGCTTAATTTTTTTTTTTT 3'].

(Pharmacia), digested with Not I and cloned into the Not I
 and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:

Pred. No.: 4,62e-151 Length: 472

Score: 156.00 Matches: 156

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 17.20% Indels: 0

DB: 1 Gaps: 0

US-10-751-736-84 (1-907) x AA460529 (1-472)

QY 566 ThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSer 585

Db 3 ACCATAGCAGTCTGGGCACCTTACTTGTATGCTTTGGTCACTTCAACAGTTTTCAGATCC 62

QY 586 ProLeuTyrlleSerProIleLysLeuLeuIleGlyValIleAlaAlaValAenMetLeu 605

Db 63 CTTCTGTAACATTTCCCCCATTAACCTGTAAATTTGGGGTCATCGCAGCAGTGAACATGCTC 122

QY 606 ThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAla 625

Db 123 ACGGGAGTCTCCAGTCCGCTGCTGGCTGTGGATGCTTTCACATTTTGGCAGCTTGCA 182

QY 626 ArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIle 645

Db 183 CGACATGCTGCTGCTGGGAGAAATGGGTGTGGTCCCATGTCAATTTTGTCCATT 242

QY 646 PheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSer 665

Db 243 TTTGCTTCAGAAATCACTGTGTTTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTCTCT 302

QY 666 VallysTyrlleSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValleIle 685

Db 303 GTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCTCGAAGTATCATTT 362

QY 686 LeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLys 705

Db 363 TTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 422

QY 706 TyrlleAlaSerProLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 721

Db 423 TATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 470

RESULT.9

BQ323949

LOCUS

DEFINITION

BQ323949

ACCESSION

BQ323949.1

KEYWORDS

EST.

SOURCE

Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

556 bp mRNA linear EST 17-MAY-2002

IL5-CT0149-011100-224-a05 C10149 Homo sapiens cDNA, mRNA sequence.

BQ323949

BQ323949

BQ323949.1

GI:20935938

Homo sapiens (human)

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.


```

REFERENCE
AUTHORS      1 (bases 1 to 556)
              Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
              Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
              Goldman,G.H., Carvalho,A.F., Mateukuma,A., Baia,G.S., Simpson,D.H.,
              Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
              O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
              Simpson,A.J.
TITLE        Shotgun sequencing of the human transcriptome with ORF expressed
              sequence tags
JOURNAL      Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE      20202663
PUBMED      10737800
COMMENT      Contact: Simpson A.J.G.
              Laboratory of Cancer Genetics
              Ludwig Institute for Cancer Research
              Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
              Brazil
              Tel: +55-11-2704922
              Fax: +55-11-2707001
              Email: asimpson@ludwig.org.br
              This sequence was derived from the FAPESP/LICR Human Cancer Genome
              Project. This entry can be seen in the following URL
              (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-CI0149-011100-224-a05&t3=2000-11-01&t4=1)
              Seq primer: puc 18 forward
              High quality sequence stop: 48.
              Location/Qualifiers
FEATURES
    source
    1..556
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /dev_stage="Adult"
        /clone_lib="CI0149"
        /note="Organ: colon_ins; Vector: puc18; Site 1: SmaI;
        Site 2: SmaI; A mini-library was made by cloning products
        derived from ORESTES PCR (U.S. Letters Patent application
        No. 196,716 - Ludwig Institute for Cancer Research)
        profiles into the pUC 18 vector. Reverse transcription of
        tissue mRNA and cDNA amplification were performed under
        low stringency conditions."
    ORIGIN
    Alignment Scores:
    Pred. No.:      5,458-151      Length:      556
    Score:          156.00         Matches:    156
    Percent Similarity: 100.00%      Conservative: 0
    Best Local Similarity: 100.00%    Mismatches: 0
    Query Match:    17.20%          Indels:     0
    DB:              5              Gaps:       0

US-10-751-736-84 (1-907) x BQ323949 (1-556)

QY      1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
DB      88 ATGGACACCTCCGGCTCGGTGCTCCTGCTCTTGGCTGCTGCTGCTGCTGCTGCGGACC 147
QY      21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
DB      148 GGGGGGCGAGCTCTCCAAAGGCTCTGGTGTGGTGTGGTGTGGGGGGGTGGCCACACACTGTCTATTGC 207
QY      41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
DB      208 GAGCCCGAGCGGAGGAGATGTGCTCAGGGTGGAGCTGCTCCGACCTGGGGGCTCTCGGAGCTG 267
QY      61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
DB      268 CCTTCCACCTCAGCGCTCTACCTCTCTACCTAGACCTCAGTATGAACAACATCAGTCAG 327
QY      81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
DB      328 CTGCTCCCGAATCCCTCGCCAGCTCTCCGCTTCTCTGGAGGAGGTACGCTTCCGGGAAAC 387
QY      101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120

```

```

Db      388 GCTCTGCATACATTCCTCAAGGAGCATTCACTGGCCTTTACAGTCTTAAAGTTCTTATG 447
QY      121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
DB      448 CTGCAGAAATAATACGTAAAGACAGTACCCACAGAAAGCTCTGCAGAAATTTCCGAAGCCTT 507
QY      141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSer 156
DB      508 CAATCCCTGGCTCTGGATGCTTAACACATCAGCTATGTGCCCCAGC 555

RESULT 10
LOCUS   BM795259
DEFINITION K-EST0077025 S21SNU520 Homo sapiens cDNA clone S21SNU520-31-C01 5',
          mRNA sequence.
ACCESSION BM795259
VERSION   BM795259
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 671)
          Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
          Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
          Kim,Y.S.
          21C Frontier Korean EST Project 2001
          Unpublished (2002)
          Contact: Kim YS
          Genome Research Center
          Korea Research Institute of Bioscience & Biotechnology
          52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
          Tel: +82-42-860-4470
          Fax: +82-42-860-4409
          Email: yongsung@mail.kribb.re.kr
          Plate: 31 row: C column: 01
          High quality sequence stop: 671.
          Location/Qualifiers
FEATURES
    source
    1..671
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="S21SNU520-31-C01"
        /sex="M"
        /tissue_type="Stomach"
        /cell_type="Floating aggregates"
        /cell_line="SNU-520"
        /lab_host="Top10F"
        /clone_lib="S21SNU520"
        /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
        Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
        bacterial alkaline phosphatase (BAP) and then decapped
        with tobacco acid pyrophosphatase (TAP). The decapped
        intact mRNA was ligated with DNA-RNA linker including EcoR
        I site by treatment of T4 RNA ligase and the first strand
        cDNA was synthesized from oligo dt-selected mRNA by
        priming with dt-tailed vector. The dt-tailed vector was
        adjusted to have about 50nt. The cDNA vector was
        circularized with E. coli DNA ligase after digestion of
        EcoRI which site is also included in vector. An RNA strand
        converted to a DNA strand by Okayama-Berg method. The
        obtained cDNA vectors were used for transformation of
        competent cells E. coli Top10F by electroporation method.
        The cDNA libraries constructed by this method are
        full-length enriched cDNA library."
    ORIGIN
    Alignment Scores:
    Pred. No.:      1,278-143      Length:      671
    Score:          149.00         Matches:    222
    Percent Similarity: 99.11%      Conservative: 0
    Best Local Similarity: 99.11%    Mismatches: 1

```



```

Db      118 TCTCTATGTAACACACATTGCCCTGTGCTTTCACCAACTGCATCTAAACGCT 59
Qy      782 ValAlaPheLeuSerPheSerLeuLeuAenLeuThrPheIleSerProGlu 799
Db      58 GTGGCTTCTTGCTCTCTCTCTTATAAACCCTACATTATCAGTCTGAA 5

RESULT 14
AI697103/c
LOCUS      539 bp mRNA linear EST 14-DEC-1999
DEFINITION tql4f08.x1 NCI CGAP Ut3 Homo sapiens cDNA clone IMAGE:2208807 3'
            similar to TR:075473 075473 ORPHAN G PROTEIN-COUPLED RECEPTOR HG38.
            ; mRNA sequence.
ACCESSION  AI697103
VERSION     AI697103.1 GI:4985003
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 539)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
JOURNAL    Contact: Robert Strausberg, Ph.D.
COMMENT     Email: cgapbs-remail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: Greg Lennon, Ph.D.
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 960 Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 405.
            Location/Qualifiers
FEATURES   source
            1..539
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:2208807"
            /tissue_type="poorly-differentiated endometrial
            adenocarcinoma, 2 pooled tumors"
            /lab_host="DH10B"
            /clone_lib="NCI-CGAP Ut3"
            /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
            Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
            Average insert size 1.45 kb. Life Technologies catalog #:
            11541-018"

ORIGIN
Alignment Scores:
Pred. No.:      3.67e-104      Length:      539
Score:          111.00      Matches:      111
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     12.24%      Indels:      0
DB:              1          Gaps:          0

US-10-751-736-84 (1-907) x AI697103 (1-539)

Qy      427 LeuAspLeuSerSerAsnLeuLeuSerPheProIleThrGlyLeuHisGlyLeuThr 446
Db      334 CTGGACCTATCGTCCAACTCTCTGCTGCTTTTCATAACTGGGTTACATGTTAACT 275
Qy      447 HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuLeuSerSerGluAsnPro 466
Db      274 CACTTAAATTAACAGGAATCATCGCTTACAGACTTGATATCATCTGAAACTTCCA 215
Qy      467 GluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGlu 486

```

```

Db      214 GAATCAAGGTATAGAAATGCTTATGCTTACCAGTGTGTCATTGGAGTGTGTGAG 155
Qy      487 AsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu 506
Db      154 AATGCTATATAGATTTCTTAATCAATGGAATAAAGTGACACAGCAGATGATGACGACCTT 95
Qy      507 HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeu 526
Db      94 CATAAGAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCCTG 35
Qy      527 LeuAspPheGluGluAspLeuLysAlaLeuHis 537
Db      34 CTTGACTTTGAGGAAGACCTGAAAGCCCTTCAT 2

RESULT 15
AY411733
LOCUS      2593 bp DNA linear GSS 16-DEC-2003
DEFINITION Pan troglodytes GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence,
            genomic survey sequence.
ACCESSION  AY411733
VERSION     AY411733.1 GI:39767701
KEYWORDS   GSS
SOURCE     Pan troglodytes (chimpanzee)
ORGANISM   Pan troglodytes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE  1 (bases 1 to 2593)
            Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
            Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
            Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
            Adams, M.D. and Cargill, M.
            Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
JOURNAL    Science 302 (5652), 1960-1963 (2003)
PUBMED     14671302
REFERENCE  2 (bases 1 to 2593)
            Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
            Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
            Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,
            Adams, M.D. and Cargill, M.
            Direct Submission
            Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT     This sequence was made by sequencing genomic exons and ordering
            them based on alignment.
FEATURES   Location/Qualifiers
            source
            1..2593
            /organism="Pan troglodytes"
            /mol_type="genomic DNA"
            /db_xref="taxon:9598"
            <1..>2593
            /gene="GPR49"
            /locus_tag="HCM4322"

ORIGIN
Alignment Scores:
Pred. No.:      6.03e-90      Length:      2593
Score:          98.00      Matches:      223
Percent Similarity: 98.24%      Conservative: 0
Best Local Similarity: 98.24%      Mismatches: 2
Query Match:     10.80%      Indels:      4
DB:              9          Gaps:          0

US-10-751-736-84 (1-907) x AY411733 (1-2593)

Qy      610 SerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArgHisGlyAla 629
Db      1697 AGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1756
Qy      630 TrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePheAlaSerGlu 649
Db      1757 TGGTGGGGAATGGGGTTGGTTGGCATGTCATGTTTTTTTGTCCATTTTGTCTTCGGA 1816

```



```
Qy      650 SerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerValIysTyrSer 669
Db      1817 TCAICTGTTTTCCTGCTTACTCTGCACGCCCTGGAGCGTGGTCTCTCTGTAATAATTTCT 1876

Qy      670 AlAlysPheGluThrLysAlaProPheSerSerLeuLysValIleIleLeuLeuCysAla 689
Db      1877 GCMAAATTTGAAACGAAAGCTCCCTTTTCTAGCCTGAAAGTAATCATTTTGTCTCTGTGCC 1936

Qy      690 LeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyrGlyAla-Se 709
Db      1937 CTGCTGGGCTTTGACCATGGCGCAGTTCCCTGCTGGGTGGCAGCAAGTATGGCAC-CTC 1995

Qy      709 rProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMetValAlaLe 729
Db      1996 CCCCTCTCGCCCTTGCCCTTTTGGGAGGCCACGACCATGGGCTACATGGTGGCTCT 2055

Qy      729 uIleLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLysLeuTyrCy 749
Db      2056 CATCTTGCTCAATTCCTTTTGTCTTCCTCATGTAGACCATTGCTTACACCAAGCTCTACTG 2115

Qy      749 aAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLysHisIleAl 769
Db      2116 CAATTTGGACAAGGGAGACCTGGAGAAATATTTGGGACTGCTCTATGGTAAAAACACATTGC 2175

Qy      769 aLeuLeuLeuPheThrAsnCysIleLeuAsn-CysProValAlaPheLeuSerPheSerS 789
Db      2176 CCTGTTGCTCTTCACCAACTGCACTCTATA-CTGCCCTGTGGCTTTCTTGCTCTCTCT 2234

Qy      789 erLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeuValValV 809
Db      2235 CTTTAATAAACCTTACATTTTATCAGTCTCTGAAGTAATTAAGTTTATCCTTCTGGTGGTAG 2294

Qy      809 alProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHisPheLysG 829
Db      2295 TCCCACTTCTCGCATGTCTCAATCCCTTCTCTACATCCTGTTCATCCTCCTCACTTTAAGG 2354

Qy      829 luAspLeuValSerLeu 834
Db      2355 AGGATCTGTGTAGGCTG 2371
```

Search completed: July 12, 2005, 12:55:36
Job time : 7011 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model
Run on: July 12, 2005, 07:36:43 ; Search time 103 Seconds
(without alignments)
3405.744 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 907
Sequence: 1 MDTSLGVLSPVLLQLAT.....PAYPTESCHLSVAFPVCL 907

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 2105692 seqs, 386760381 residues

Word size : 0
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

- Database : A_Geneseq_16Dec04.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	907	100.0	907	2	Aaw93889 Human HG3
2	907	100.0	907	3	Aay90682 Human G p
3	907	100.0	907	6	Abp81968 Human G p
4	907	100.0	907	6	Abc06467 Human G-p
5	907	100.0	907	6	Adc22783 Human G p
6	907	100.0	907	7	Adc59150 Human Pro
7	907	100.0	907	7	Adc59153 Human Pro
8	907	100.0	907	7	Adg42628 Human G p
9	907	100.0	907	7	Adh14256 Human HG3
10	907	100.0	907	7	Adn40013 Cancer/an
11	907	100.0	907	7	Adn39531 Cancer/an
12	907	100.0	907	7	Adn39628 Cancer/an
13	907	100.0	907	8	Ado29408 Human GPC
14	907	100.0	907	8	Adg80369 G protein
15	907	100.0	907	8	Adf67868 Human HG3
16	806	88.9	907	3	Aay90687 Human mut
17	806	88.9	907	7	Adc22797 Human G p
18	806	88.9	907	7	Adh14270 Mutated h
19	806	88.9	1145	7	Adf70480 Orphan re
20	716	78.9	907	7	Adg42629 Human G p
21	689	76.0	692	2	Aaw93890 Human HG3
22	621	68.5	883	7	Adb80464 Ovarian c
23	621	68.5	883	7	Adn40012 Cancer/an
24	621	68.5	883	7	Adn39166 Cancer/an
25	592	65.3	693	2	Aay42169 Human LGR

26	202	22.3	202	2	AAV53573	Aay53573 Human gon
27	197	21.7	282	5	ABP42000	Abp42000 Human ova
28	51	5.6	907	8	ADG75449	Adg75449 Mouse orp
29	51	5.6	907	8	ADO29409	Ado29409 Mouse GPC
30	43	4.7	43	2	AAW93892	Aaw93892 Human HG3
31	43	4.7	43	2	AAW93894	Aaw93894 Human HG3
32	43	4.7	43	2	AAW93893	Aaw93893 Human HG3
33	43	4.7	43	2	AAW93895	Aaw93895 Human HG3
34	43	4.7	43	2	AAW93896	Aaw93896 Human HG3
35	43	4.7	43	2	AAW93891	Aaw93891 Human HG3
36	20	2.2	20	6	ABP82972	Abp82972 G protein
37	20	2.2	20	6	ABP82976	Abp82976 G protein
38	18	2.0	91	2	AAV53577	Aay53577 Human gon
39	18	2.0	244	8	AQG66490	Adq66490 Novel hum
40	18	2.0	266	3	AAAB41583	Aab41583 Human ORP
41	18	2.0	736	5	AAG66140	Aag66140 Human LGR
42	18	2.0	736	7	ADN02246	Adn02246 Human par
43	18	2.0	736	8	ADK19411	Adk19411 Human LGR
44	18	2.0	794	5	ADH50810	Adh50810 Human G-p
45	18	2.0	823	5	ADH50809	Adh50809 Human G-p

ALIGNMENTS

RESULT 1
AAW93889
ID AAW93889 standard; protein; 907 AA.

AC AAW93889;

DT 25-JUN-1999 (first entry)

DE Human HG38 protein.

KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain;
endocrine system; skeletal muscle; spinal cord; placenta; development;
KW receptor activity modulator.

OS Homo-sapiens.

PN WO9915660-A1-

XX 01-APR-1999.

XX 24-SEP-1998; 98WO-US019979.

XX 24-SEP-1997; 97US-0059863P.

XX (MERI) MERCK & CO INC.

XX Liu Q, Bailey WJ, McDonald TP;

XX WPI; 1999-254711/21.

XX N-PSDB; AAX23980.

XX Human G-protein coupled glycoprotein hormone receptor HG38.

XX Claim 1a; Fig 2; 74pp; English.

CC This invention describes a novel human G-protein coupled glycoprotein hormone receptor, HG38. Glycoprotein hormone receptors are important in the endocrine system and HG38 may be involved in development and function of the skeletal muscle, spinal cord, placenta and to a lesser extent, the brain. The transgenic animal may be useful for studying tissue and temporal specific expression or activity of the HG38 receptor, as well as for studying the ability of a variety of compounds to act as modulators of HG38 receptor activity

SQ Sequence 907 AA;

Query Match 100.0%; Score 907; DB 2; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

[illegible]


```
QY 181 AFRSLSALQAMTLANKIHHPDYAFGNLSLVVLHLNHNRIHSLGKKCFDGLHSLETLD 240
D 181 AFRSLSALQAMTLANKIHHPDYAFGNLSLVVLHLNHNRIHSLGKKCFDGLHSLETLD 240
QY 241 LNYNLDLDFPTAIRLTLNKLGHFHSNNIRISPEKAFVGNPSLTIHFYDNPQFVGRSA 300
D 241 LNYNLDLDFPTAIRLTLNKLGHFHSNNIRISPEKAFVGNPSLTIHFYDNPQFVGRSA 300
QY 301 FQHLPELRTLTLNGASQITTEPDLTGTLNLSLTLTGAISSLPQVTCNQLPNLQVLDLS 360
D 301 FQHLPELRTLTLNGASQITTEPDLTGTLNLSLTLTGAISSLPQVTCNQLPNLQVLDLS 360
QY 361 YNLEDLPSFVCQKLOKIDLRHNEIYEIKVDTFQOLLSLRLSLNLANWKIAIHPNAPST 420
D 361 YNLEDLPSFVCQKLOKIDLRHNEIYEIKVDTFQOLLSLRLSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKDLNLSNLSRPFITGLHGTUHLKLTGNHALQSLISSENPPKLVTEMPPAYOCC 480
D 421 LPSLIKDLNLSNLSRPFITGLHGTUHLKLTGNHALQSLISSENPPKLVTEMPPAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
D 481 AFGVCENAYKISNOWNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPPKPEHLIDGMLIRIGVWTVIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
D 541 CSPSPGPPKPEHLIDGMLIRIGVWTVIAVLTALTCNALVTSTVFRSPLYSPIKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSPFARHGAWENGVCVIGFLSTFASSESVFLLTLAAL 660
D 601 AVNMLTGVSVAVLGADFTGSPFARHGAWENGVCVIGFLSTFASSESVFLLTLAAL 660
QY 661 ERGSFVKYSAKFETKAPFSSLVKVIILLCALLALTAAPVLLGGSKYKASPLCLPLPFGEP 720
D 661 ERGSFVKYSAKFETKAPFSSLVKVIILLCALLALTAAPVLLGGSKYKASPLCLPLPFGEP 720
QY 721 STMGVMVALIILNLSLCFLMTIATYKLYCNLDKGDLENWDCSMVKHIALLLFNCLINC 780
D 721 STMGVMVALIILNLSLCFLMTIATYKLYCNLDKGDLENWDCSMVKHIALLLFNCLINC 780
QY 781 PVAFLSFSSNLNLFISBEVIFKILLVVVLPACLNPLLYLILFPHFKEDLVLSRKQYV 840
D 781 PVAFLSFSSNLNLFISBEVIFKILLVVVLPACLNPLLYLILFPHFKEDLVLSRKQYV 840
QY 841 WTRSKHPSLMSINDVDEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
D 841 WTRSKHPSLMSINDVDEKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVPTESCHLSS 900
QY 901 VAFVPCPL 907
D 901 VAFVPCPL 907
```

RESULT 3

ABP81968
ID ABP81968 standard; protein; 907 AA.

AC ABP81968;

DT 04-MAR-2003 (first entry)

DE Human G protein-coupled receptor GPR49 protein SEQ ID NO:422.

DE KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;

ulcer.

OS Homo sapiens.

XX WO200261087-A2.

XX 08-AUG-2002.

XX 19-DEC-2001; 2001WO-US050107.

XX 19-DEC-2000; 2000US-0257144P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Burmer GC, Roush CL, Brown JP;

XX WPI; 2003-046718/04.

XX N-PSDB; ABZ42816.

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode GPCR proteins given in ABP81675 to ABP82018, which are used in the exemplification of the present invention

Sequence 907 AA;

Query Match 100.0%; Score 907; DB 6; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MDTSLRGLVLLSLPVLLQLATGSSPSRSGVTLRGCTHCEPDGMRMLLRVDCSDGLSEL 60
D 1 MDTSLRGLVLLSLPVLLQLATGSSPSRSGVTLRGCTHCEPDGMRMLLRVDCSDGLSEL 60
QY 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSRLEELRLAGNALTYIPKGAFTGLYSKVLV 120
D 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSRLEELRLAGNALTYIPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSYVPPSCFSGLSLRLHLDNALTETIPVQ 180
D 121 LQNNQLRHVPTEALQNLRSLSQSLRLDANHSYVPPSCFSGLSLRLHLDNALTETIPVQ 180
QY 181 AFRSLSALQAMTLANKIHHPDYAFGNLSLVVLHLNHNRIHSLGKKCFDGLHSLETLD 240
D 181 AFRSLSALQAMTLANKIHHPDYAFGNLSLVVLHLNHNRIHSLGKKCFDGLHSLETLD 240
```


QY 241 LNNYNDLDEFPPTAIRTLNLKELGPHSNIRISPEKAFVGNPSLTIHFYDNPDIQVGRSA 300
DB 241 LNNYNDLDEFPPTAIRTLNLKELGPHSNIRISPEKAFVGNPSLTIHFYDNPDIQVGRSA 300
QY 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLSPQTCVQCNLPNLQVLDLS 360
DB 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLSPQTCVQCNLPNLQVLDLS 360
QY 361 YNLLEDLPFSVCOKLQKIDLRHNEIYEIKVDYTFQQLLSRLSLNLANWKIAIHPNAPST 420
DB 361 YNLLEDLPFSVCOKLQKIDLRHNEIYEIKVDYTFQQLLSRLSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLLSPFITGLHGTHTLKTGNHALQSLISENPFELKVIEMPVAYOCC 480
DB 421 LPSLIKLDLSSNLLSPFITGLHGTHTLKTGNHALQSLISENPFELKVIEMPVAYOCC 480
QY 481 AFGVCENAYKISNOWNKGDNSMDLHKDDAGMFOAQDERDLEDPLDFEEDLKALHSVQ 540
DB 481 AFGVCENAYKISNOWNKGDNSMDLHKDDAGMFOAQDERDLEDPLDFEEDLKALHSVQ 540
QY 541 CSPSPGPKPEHLLDGLHRIIGWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB 541 CSPSPGPKPEHLLDGLHRIIGWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSVAVLGADFTGSPARHGAWENGVCVIGFLSIFASESSVFLTLAAL 660
DB 601 AVNMLTGVSVAVLGADFTGSPARHGAWENGVCVIGFLSIFASESSVFLTLAAL 660
QY 661 ERGSVKYSKAFETKAPFSSLKVIILLCALLATMAAVPLLGSSKYGASPLCLPFGEP 720
DB 661 ERGSVKYSKAFETKAPFSSLKVIILLCALLATMAAVPLLGSSKYGASPLCLPFGEP 720
QY 721 STMGVWALILNLSLCFLMTIATYKLYCNLDKGDLENWDCSWKHVHALLFTNCILNC 780
DB 721 STMGVWALILNLSLCFLMTIATYKLYCNLDKGDLENWDCSWKHVHALLFTNCILNC 780
QY 781 PVAFLSFSSLLNLTPISEVTKFILLVVVPLPACLNPLLYILFNPHKEDLVSLRKQTYV 840
DB 781 PVAFLSFSSLLNLTPISEVTKFILLVVVPLPACLNPLLYILFNPHKEDLVSLRKQTYV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTTESCHLSS 900
DB 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTTESCHLSS 900
QY 901 VAFVPCL 907
DB 901 VAFVPCL 907

RESULT 4
ABO06467
ID ABO06467 standard; protein; 907 AA.
XX
AC ABO06467;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human G-protein coupled receptor HG38.
XX
KW G-protein coupled receptor; GPCR; caudate nucleus related disorder;
KW neoplastic disorder; leukaemia; breast cancer; immune disorder; AIDS;
KW rheumatoid arthritis; neuronal disorder; Alzheimer's disease; receptor;
KW Parkinson's disease; respiratory disorder; bronchopulmonary disease;
KW pancreatic disease; ovarian disorder; ovarian carcinoma; colon disease;
KW Meig's syndrome; breast disease; thalamus related disorder; gene therapy;
KW amygdala related disorder; corpus callosum related disorder;
KW hippocampus related disorder.
XX
OS Homo sapiens.
XX
PN US2003027323-A1.
XX

PD 06-FEB-2003.
XX
PF 26-SEP-2001; 2001US-00965536.
XX
PR 27-SEP-2000; 2000US-0235832P.
PR 16-JAN-2001; 2001US-0261781P.
PR 19-JUL-2001; 2001US-030605P.
PR 03-AUG-2001; 2001US-0310436P.
XX
PA (FEDE/) FEDER J N.
PA (MINT/) MINTIER G.
PA (RAMA/) RAMANATHAN C S.
PA (HAWK/) HAWKEN D R.
XX
PI Feder JN, Mintier G, Ramanathan CS, Hawken DR;
XX WPI; 2003-466147/44.

New G-protein coupled receptor polypeptides, designated as HGPRBMY5, useful for preventing, treating or ameliorating a medical condition related to the colon, breast, ovaries or immune system.

Disclosure; Fig 8; 93pp; English.

The invention relates to an isolated HGPRBMY5 polypeptide. The polypeptides, polynucleotides and methods are useful for preventing, treating or ameliorating a medical condition such as a neoplastic disorder e.g. leukaemia and breast cancer; immune disorder e.g. AIDS and rheumatoid arthritis; neuronal disorder e.g. Alzheimer's disease and Parkinson's disease; respiratory disorder e.g. bronchopulmonary disease and pancreatic disease; ovarian disorder e.g. ovarian carcinoma and Meig's syndrome; colon disease; breast disease; thalamus related disorder; amygdala related disorder; corpus callosum related disorder; caudate nucleus related disorder; hippocampus related disorder by administering the GPCR polypeptide or its homologue. The present sequence represents the amino acid sequence of a G-protein coupled receptor used to show homology with the human G-protein coupled receptor, HGPRBMY

Sequence 907 AA;

Query Match 100.0%; Score 907; DB 6; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLRGCTHCHCEPDGRMLLRVDCSDLGSEL 60
DB 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLRGCTHCHCEPDGRMLLRVDCSDLGSEL 60
QY 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRELELRAGNALTYPKGAFTGLYSKVLV 120
DB 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSLRELELRAGNALTYPKGAFTGLYSKVLV 120
QY 121 LQNNQLRHVPTEALQNLRSQSLRLDANHSIYVPPSCPSGLHSRLHLDNALTETIPIVQ 180
DB 121 LQNNQLRHVPTEALQNLRSQSLRLDANHSIYVPPSCPSGLHSRLHLDNALTETIPIVQ 180
QY 181 AFRSLSALQAMTALNKITHIHPDYAFGNLSLVVLHNNRHISLGKCKFQGLHSLETLD 240
DB 181 AFRSLSALQAMTALNKITHIHPDYAFGNLSLVVLHNNRHISLGKCKFQGLHSLETLD 240
QY 241 LNNYNDLDEFPPTAIRTLNLKELGPHSNIRISPEKAFVGNPSLTIHFYDNPDIQVGRSA 300
DB 241 LNNYNDLDEFPPTAIRTLNLKELGPHSNIRISPEKAFVGNPSLTIHFYDNPDIQVGRSA 300
QY 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLSPQTCVQCNLPNLQVLDLS 360
DB 301 FOHLPELRTLTNGASQITEFPDLTGTANLESRLTGAQISLSPQTCVQCNLPNLQVLDLS 360
QY 361 YNLLEDLPFSVCOKLQKIDLRHNEIYEIKVDYTFQQLLSRLSLNLANWKIAIHPNAPST 420
DB 361 YNLLEDLPFSVCOKLQKIDLRHNEIYEIKVDYTFQQLLSRLSLNLANWKIAIHPNAPST 420
QY 421 LPSLIKLDLSSNLLSPFITGLHGTHTLKTGNHALQSLISENPFELKVIEMPVAYOCC 480

Db 421 LPSLKLSSNLLSPFTGLHGLTHLKTGNHALQSLSSNPELKVEMPAYQCC 480
Qy 481 AFGVCENAYKISQWKNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQWKNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCBHLDDGMLIRIGVMTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGMLIRIGVMTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Qy 601 AVNMLTGSSAVLAGVDAFTGSGFARHGAWNGVCHVIGLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGSSAVLAGVDAFTGSGFARHGAWNGVCHVIGLSIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSAKETKAPSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPFGEP 720
Db 661 ERGFSVKYSAKETKAPSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPFGEP 720
Qy 721 STMGYVVALILNLSLCLMFTIATYTKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVVALILNLSLCLMFTIATYTKLYCNLDKGLNIWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFPSRSLNLTISPEVIFKILLVVPPLPACLNPLLYILPNPHKEDVLSLRQTYV 840
Db 781 PVAFPSRSLNLTISPEVIFKILLVVPPLPACLNPLLYILPNPHKEDVLSLRQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTECHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFSSITVDLPSSVPSPAYPVTECHLSS 900
Qy 901 VAFVPCCL 907
Db 901 VAFVPCCL 907
RESULT 5
ID ADC22783 standard; protein; 907 AA.
XX ADC22783;
XX 18-DEC-2003 (first entry)
XX Human G protein-coupled receptor (GPCR) polypeptide #39.
XX Human; G protein-coupled receptor; GPCR; transmembrane-6 region; TM6;
XX intracellular-3 region; IC3; receptor.
XX Homo sapiens.
XX US6555339-B1.
XX 29-APR-2003.
XX 13-OCT-1998; 98US-00170496.
XX 14-APR-1997; 97US-00839449.
XX 14-APR-1998; 98US-00060188.
XX 26-JUN-1998; 98US-0090783P.
XX 07-AUG-1998; 98US-0095677P.
XX (AREN-) ARENA PHARM INC.
XX Liaw CW, Behan DP, Chalmers DT;
XX WPI; 2003-742861/70.
XX N-PSDB; ADC22782.
XX Creating a constitutively active version of an endogenous human G protein
XX coupled receptor (GPCR) comprises substituting a specific amino acid in
XX the transmembrane-6 region with a different amino acid, and testing for
XX constitutive activity.

XX Example 1; SEQ ID NO 264; 221pp; English.
XX The invention relates to a method for treating a non-endogenous,
XX constitutively active version of an endogenous human G protein-coupled
XX receptor (GPCR) that has a transmembrane-6 (TM6) region and an
XX intracellular-3 (IC3) region, by substituting a specific amino acid in
XX the TM6 region with a different amino acid, and testing for constitutive
XX activity. The method is useful for creating a constitutively active
XX version of an endogenous human GPCR that comprises a transmembrane 6
XX region and an intracellular loop 3 region. The altered human GPCR
XX polypeptides are useful for screening test compounds for identification
XX of inverse agonists or partial agonists of GPCR polypeptides, which may
XX have therapeutic uses. The altered GPCRs may also be used in vivo or in
XX vitro in biological research. A nucleic acid encoding the altered GPCR
XX may be used to create a transgenic animal expressing the altered GPCR.
XX The method allows screening for compounds that modulate the activity of a
XX human G protein-coupled receptor without the need for provision of a
XX ligand for the receptor. This is particularly useful in allowing
XX screening of compounds against orphan receptors for which no ligand is
XX currently known. This sequence represents a human GPCR polypeptide of the
XX invention.
XX Sequence 907 AA;
Qy Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDTSELGVLLSLPVLQLLATGSSPSRSGVLLRGCPDTHCEPDGSMLLRVDCSDGLSEL 60
Db 1 MDTSELGVLLSLPVLQLLATGSSPSRSGVLLRGCPDTHCEPDGSMLLRVDCSDGLSEL 60
Qy 61 PSNLSVFTSYLDLSNMNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLVLM 120
Db 61 PSNLSVFTSYLDLSNMNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLVLM 120
Qy 121 LQNNQLRHVPTEALQNLRSLSQSLRDANHSIYVPPSCFSGLSLHRLMDDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSQSLRDANHSIYVPPSCFSGLSLHRLMDDNALTEIPVQ 180
Qy 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLHNNRHSHSGKCFDGLHSLTD 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVLHNNRHSHSGKCFDGLHSLTD 240
Qy 241 LNYNMLDEFPPTAIRTLNKLKELGPHSNIRISPEKAFVGNPSLITIHFDNPIQFVGRSA 300
Db 241 LNYNMLDEFPPTAIRTLNKLKELGPHSNIRISPEKAFVGNPSLITIHFDNPIQFVGRSA 300
Qy 301 FOHLPELTLTLNGASQITEFPDLTGTLNLESLSLTGAQISLPTQVCNQLPNLQVLDLS 360
Db 301 FOHLPELTLTLNGASQITEFPDLTGTLNLESLSLTGAQISLPTQVCNQLPNLQVLDLS 360
Qy 361 YNLLLEDLPSFVSCQKLOKIDLRHNEIYEIKVDFTFOQLLSLSLNLAWNKIAIHNAPST 420
Db 361 YNLLLEDLPSFVSCQKLOKIDLRHNEIYEIKVDFTFOQLLSLSLNLAWNKIAIHNAPST 420
Qy 421 LPSLKLDSNLLSSPFTGLHGLTHLKTGNHALQSLSSNPELKVEMPAYQCC 480
Db 421 LPSLKLDSNLLSSPFTGLHGLTHLKTGNHALQSLSSNPELKVEMPAYQCC 480
Qy 481 AFGVCENAYKISQWKNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQWKNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCBHLDDGMLIRIGVMTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Db 541 CSPSPGPKPCBHLDDGMLIRIGVMTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600
Qy 601 AVNMLTGSSAVLAGVDAFTGSGFARHGAWNGVCHVIGLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGSSAVLAGVDAFTGSGFARHGAWNGVCHVIGLSIFASESSVFLTLAAL 660

Qy 661 ERGFSVKYSAKFTKAPSSSLKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEF 720
 Db |||||||
 Qy 661 ERGFSVKYSAKFTKAPSSSLKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEF 720
 Db |||||||
 Qy 721 STMGYVALIILNSLCFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
 Db |||||||
 Qy 721 STMGYVALIILNSLCFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
 Db |||||||
 Qy 781 PVAFSLFSSNLNLTPISEVVKIFILLVVVPLPACLNPLLYILFNPHKEDLVSLRKQYV 840
 Db |||||||
 Qy 781 PVAFSLFSSNLNLTPISEVVKIFILLVVVPLPACLNPLLYILFNPHKEDLVSLRKQYV 840
 Db |||||||
 Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFITSSSITVDLPSSVPSPAYPVTSSCHLSS 900
 Db |||||||
 Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFITSSSITVDLPSSVPSPAYPVTSSCHLSS 900
 Db |||||||
 Qy 901 VAFVPECL 907
 Db |||||||
 Qy 901 VAFVPECL 907
 Db |||||||

RESULT 6
 ADE59150
 ID ADE59150 standard; protein; 907 AA.
 XX
 AC ADE59150;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein O75473, SRQ ID NO 5041.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 XX (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 XX WPI; 2003-268312/26.
 DR GENBANK; O75473.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX
 PS Claim 1; Page; 1017pp; English.
 XX

The invention discloses a composition comprising two or more isolated rat
 or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a human protein (shown in table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 907 AA;

Query Match 100.0%; Score 907; DB 7; Length 907;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSRGLGVLLSLPVLLQLATGCGSSPRSGVLLRGCPTHCHCEPDGRMLLRVDCSDLGSEL 60
 Db |||||||
 Qy 1 MDTSRGLGVLLSLPVLLQLATGCGSSPRSGVLLRGCPTHCHCEPDGRMLLRVDCSDLGSEL 60
 Db |||||||
 Qy 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSRLRFLBELRLAGNALTYIPKGAFTGLYSKVLVLM 120
 Db |||||||
 Qy 61 PSNLSVFTSYLDLSMNNISQLLPNPLPSRLRFLBELRLAGNALTYIPKGAFTGLYSKVLVLM 120
 Db |||||||
 Qy 121 LQNNQLRHVPTEALQNLRLSLQSLRLDANHISYVPPSCFSGHLHSRLHLDLNNALTEIPVQ 180
 Db |||||||
 Qy 121 LQNNQLRHVPTEALQNLRLSLQSLRLDANHISYVPPSCFSGHLHSRLHLDLNNALTEIPVQ 180
 Db |||||||
 Qy 181 AFRSLSALOAMTLALNKIHHPDYAFGNLSLVVHLHNNRIHSLGKCKCPGLHSLEFLD 240
 Db |||||||
 Qy 181 AFRSLSALOAMTLALNKIHHPDYAFGNLSLVVHLHNNRIHSLGKCKCPGLHSLEFLD 240
 Db |||||||
 Qy 241 LNYNNLDEFFTAIRTLNKLKELGPHSNINIRSIPEKAFVGNPSLTIHFYDNPIDQVGRSA 300
 Db |||||||
 Qy 241 LNYNNLDEFFTAIRTLNKLKELGPHSNINIRSIPEKAFVGNPSLTIHFYDNPIDQVGRSA 300
 Db |||||||
 Qy 301 FOHLPELRTLTLNGASQITTEPDLTGTALESLETLTGAQISSLPQTVCNQLPNIQVLDLS 360
 Db |||||||
 Qy 301 FOHLPELRTLTLNGASQITTEPDLTGTALESLETLTGAQISSLPQTVCNQLPNIQVLDLS 360
 Db |||||||
 Qy 361 YNLLLEDLPFSVCOKLQKIDLRHNEIYEIKVDTFQQLLSLSLSLANNKIAIHPNAPST 420
 Db |||||||
 Qy 361 YNLLLEDLPFSVCOKLQKIDLRHNEIYEIKVDTFQQLLSLSLSLANNKIAIHPNAPST 420
 Db |||||||
 Qy 421 LPSLIKLDLSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENFPPELKVTEMPIYAYOCC 480
 Db |||||||
 Qy 421 LPSLIKLDLSNLLSSPFIITGLHGLTHLKTGNHALQSLISSENFPPELKVTEMPIYAYOCC 480
 Db |||||||
 Qy 481 AFGVCENAYKITSNOWNKGDNSMDDLHKKODAGMFOAODERDLEFLDFFEDLKALHSVQ 540
 Db |||||||
 Qy 481 AFGVCENAYKITSNOWNKGDNSMDDLHKKODAGMFOAODERDLEFLDFFEDLKALHSVQ 540
 Db |||||||
 Qy 541 CSPSPGPKPCHEHLLDGLIRIGVWTTAVLATCNALVTSTVFRSPLYSIPKLLIGVIA 600
 Db |||||||
 Qy 541 CSPSPGPKPCHEHLLDGLIRIGVWTTAVLATCNALVTSTVFRSPLYSIPKLLIGVIA 600
 Db |||||||
 Qy 601 AVNMLTGVSAAVLAVGDAFTFGSPARHAGWENGVGCHVIGFLSIFASESSVFLTLAAL 660
 Db |||||||
 Qy 601 AVNMLTGVSAAVLAVGDAFTFGSPARHAGWENGVGCHVIGFLSIFASESSVFLTLAAL 660
 Db |||||||
 Qy 661 ERGFSVKYSAKFTKAPSSSLKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEF 720
 Db |||||||
 Qy 661 ERGFSVKYSAKFTKAPSSSLKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPFGEF 720
 Db |||||||
 Qy 721 STMGYVALIILNSLCFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
 Db |||||||
 Qy 721 STMGYVALIILNSLCFLMWTIATYKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780
 Db |||||||

QY 781 PVAFISFSSLLINLTFTSPVIFKILLVVVPLACINPLIYILFNPHFKEDLVSLRKQTVY 840
 Db 781 PVAFISFSSLLINLTFTSPVIFKILLVVVPLACINPLIYILFNPHFKEDLVSLRKQTVY 840
 QY 841 WTRSKHPSLMSINSDDEKQSCDSQALVTFSSSITYDLPPSSVPSPAYPVTESCHLSS 900
 Db 841 WTRSKHPSLMSINSDDEKQSCDSQALVTFSSSITYDLPPSSVPSPAYPVTESCHLSS 900

QY 901 VAFVPCFL 907
 Db 901 VAFVPCFL 907

RESULT 7

ADE59153
 ID ADE59153 standard; protein; 907 AA.

XX AC ADE59153;

DT 29-JAN-2004 (first entry)

XX Human Protein O75473, SEQ ID NO 5044.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; Chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-034382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

DR GENBANK; O75473.

PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.

PS Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (Chung), chronic constriction

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a human protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed CC specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 907 AA;

Query Match 100.0%; Score 907; DB 7; Length 907;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLGVLLSLPVLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLVDCSDGLSEL 60

Db 1 MDTSLGVLLSLPVLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLVDCSDGLSEL 60

QY 61 PSNLSVFYSYLDLSMNNISQLLPNPLPSLRPEELRAGNALTYPKGAFGTGLYSKVLV 120

Db 61 PSNLSVFYSYLDLSMNNISQLLPNPLPSLRPEELRAGNALTYPKGAFGTGLYSKVLV 120

QY 121 LQNNQLRHVPTEALQNLRSLSQSLRDANHSYVPPSCFSGLSRLHRLWDDNALTEIPVQ 180

Db 121 LQNNQLRHVPTEALQNLRSLSQSLRDANHSYVPPSCFSGLSRLHRLWDDNALTEIPVQ 180

QY 181 AFRSLALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLTLD 240

Db 181 AFRSLALQAMTLALNKIHHIPDYAFGNLSLVVHLHNNRIHSLGKKCFDGLHSLTLD 240

QY 241 LNNYNNLDEFPPTAIRTLNKLGEHNSNIRTSPEKAFVGNPSLTIHFVDNPIQFVGSA 300

Db 241 LNNYNNLDEFPPTAIRTLNKLGEHNSNIRTSPEKAFVGNPSLTIHFVDNPIQFVGSA 300

QY 301 FOHLPELRTLINGASQITEFPDLTGTANLESRLTGAQISSLPQTCNQLNQLVDLS 360

Db 301 FOHLPELRTLINGASQITEFPDLTGTANLESRLTGAQISSLPQTCNQLNQLVDLS 360

QY 361 YNLEEDLPFSVQCQKIDLRHNEIYEIKVDTFQOLLSLRSLNLANWKIAIHPNAPST 420

Db 361 YNLEEDLPFSVQCQKIDLRHNEIYEIKVDTFQOLLSLRSLNLANWKIAIHPNAPST 420

QY 421 LPSLIKLDLSSNLLSSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPVAQCC 480

Db 421 LPSLIKLDLSSNLLSSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPVAQCC 480

QY 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOQDERDLEDFLLDPEEDLKALHSVQ 540

Db 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOQDERDLEDFLLDPEEDLKALHSVQ 540

QY 541 CSPSPGPKPCHEHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Db 541 CSPSPGPKPCHEHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

QY 601 AVNMLTGSSAVLAGVDATFGSFARHGAWNGVCHVIGLSIFASESSVFLTLAAL 660

Db 601 AVNMLTGSSAVLAGVDATFGSFARHGAWNGVCHVIGLSIFASESSVFLTLAAL 660

QY 661 ERGFSVKYSAPKETKAPFSSLVKVIILLCALLATMAAVPLLGSKYKYGASPLCLPFPGBP 720

Db 661 ERGFSVKYSAPKETKAPFSSLVKVIILLCALLATMAAVPLLGSKYKYGASPLCLPFPGBP 720

QY 721 STMGYMVALILLNSLCFLMWTIATYTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780

Db 721 STMGYMVALILLNSLCFLMWTIATYTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780

QY 781 PVAFISFSSLLINLTFTSPVIFKILLVVVPLACINPLIYILFNPHFKEDLVSLRKQTVY 840

Db 781 PVAFISFSSLLINLTFTSPVIFKILLVVVPLACINPLIYILFNPHFKEDLVSLRKQTVY 840

QY 841 WTRSKHPSLMSINSDDEKQSCDSQALVTFSSSITYDLPPSSVPSPAYPVTESCHLSS 900

Db 841 WTRSKHPSLMSINSDDEKQSCDSQALVTFSSSITYDLPPSSVPSPAYPVTESCHLSS 900


```
QY 901 VAFVPC 907
Db 901 VAFVPC 907

RESULT 8
ID ADG42628
XX ADG42628 standard; protein; 907 AA.
AC ADG42628;
XX
DT 26-FEB-2004 (first entry)
XX
DE Human G protein-coupled receptor 49.
XX
KW cytostatic; gene therapy; NOVX-agonist; NOVX-antagonist; pharmaceutical;
KW NOVX-associated disorder; cancer; human; G protein coupled receptor 49.
XX
OS Homo sapiens.
XX
PN US2003204052-A1.
XX
XX 30-OCT-2003.
XX
XX 04-OCT-2001; 2001US-00970944.
XX
XX 04-OCT-2000; 2000US-0237862P.
XX
PA (HERR/) HERRMANN J L.
PA (RAST/) RASTELLI L.
PA (SHIM/) SHIMKETS R A.
XX
PI Herrmann JL, Rastelli L, Shinkets RA;
XX
DR WPI; 2003-900673/82.
XX
XX
PT New NOVX gene or NOVX-specific antibody, useful for preparing a
PT composition for treating or preventing a NOVX-associated disorder, e.g.,
PT cancer.
XX
XX Disclosure; SEQ ID NO 26; 118pp; English.
XX
CC The invention describes a new isolated polypeptide comprising: a
CC polypeptide or its mature form comprising a sequence not given in the
CC specification; or a variant of (A), where one or more amino acid residues
CC in the variant differs in no more than 15% from the amino acid sequence
CC of the mature form. The pharmaceutical composition may be administered
CC via oral, transdermal, rectal or parenteral route. The polypeptide,
CC nucleic acid or antibody is useful for preparing a composition for
CC treating or preventing a NOVX-associated disorder, e.g., cancer. This is
CC the amino acid sequence of a transmembrane receptor homologue used in a
CC comparison with the novel human proteins of the invention.
XX
SQ Sequence 907 AA;

Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60

QY 61 PSLNSVFTSYDLSWNNISQLLPNLPFLFLELRAGNALTYTPKGAFTGLYSKVLV 120
Db 61 PSLNSVFTSYDLSWNNISQLLPNLPFLFLELRAGNALTYTPKGAFTGLYSKVLV 120

QY 121 LQNNQLRHVPTEALQNLSRLSDANHSIVVPSCFSGLSRLHRLMDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLSRLSDANHSIVVPSCFSGLSRLHRLMDNALTEIPVQ 180

QY 181 AFRSLSALQAMTALANKIHIPDYAFGNLSLVVLHLHNNRHHISLGKCFDGLHSLTLD 240
Db 181 AFRSLSALQAMTALANKIHIPDYAFGNLSLVVLHLHNNRHHISLGKCFDGLHSLTLD 240

181 AFRSLSALQAMTALANKIHIPDYAFGNLSLVVLHLHNNRHHISLGKCFDGLHSLTLD 240
241 LNNYNDLDEFPTAIRTLNKLKELGFHSHNNIRSIPEKAFVGNPSLTIHFDYDNPVGVGSA 300
241 LNNYNDLDEFPTAIRTLNKLKELGFHSHNNIRSIPEKAFVGNPSLTIHFDYDNPVGVGSA 300
301 FOHLPELRTLTNGASQITTEPPDLTGTAANLESLLTGAQISSLPQTVCNQLPNLQVLDLS 360
301 FOHLPELRTLTNGASQITTEPPDLTGTAANLESLLTGAQISSLPQTVCNQLPNLQVLDLS 360
361 YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDVTFQQLLSLRSLNLANMKIAIHPNAPST 420
361 YNLLEDLPFSVCQKQKIDLRHNEIYEIKVDVTFQQLLSLRSLNLANMKIAIHPNAPST 420
421 LPSLIKLDLSSNLSSPITGLHGLTHLKTGNHALQSLISSENFPELKVTEMPEYVOC 480
421 LPSLIKLDLSSNLSSPITGLHGLTHLKTGNHALQSLISSENFPELKVTEMPEYVOC 480
481 AFGVCENAYKISNOWNKGDSSMDLHKDAGMFOQDERDLEDFLLDFFEDLKLHHSVQ 540
481 AFGVCENAYKISNOWNKGDSSMDLHKDAGMFOQDERDLEDFLLDFFEDLKLHHSVQ 540
541 CSPSPGPFKPCHEHLLDGWLIIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
541 CSPSPGPFKPCHEHLLDGWLIIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
601 AVNMLTGVSASVLAGVDAFTFGSFAHGAHWENGVGCHVIGFLSIFASESVFLLTAAAL 660
601 AVNMLTGVSASVLAGVDAFTFGSFAHGAHWENGVGCHVIGFLSIFASESVFLLTAAAL 660
661 ERGFSVKYSAKFETKAPFSSKLVIIICALLALTAALVPLLGSGKYGASPLCLPLFGEP 720
661 ERGFSVKYSAKFETKAPFSSKLVIIICALLALTAALVPLLGSGKYGASPLCLPLFGEP 720
721 STMGYVVALIILNSLCFLMMTIAVTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
721 STMGYVVALIILNSLCFLMMTIAVTKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC 780
781 PVAFLSFSSILNLPFISPEVIKFIILVVVPLACINPLLYILFNPHKEDLVSLRKQTYV 840
781 PVAFLSFSSILNLPFISPEVIKFIILVVVPLACINPLLYILFNPHKEDLVSLRKQTYV 840
841 WTRSKHPSLMSINDSDVEKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVPTESCHLSS 900
841 WTRSKHPSLMSINDSDVEKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVPTESCHLSS 900
901 VAFVPC 907
901 VAFVPC 907

RESULT 9
ADH14256
ID ADH14256 standard; protein; 907 AA.
XX
XX ADH14256;
XX
DT 11-MAR-2004 (first entry)
XX
DE Human HG38.
XX
KW human; non-endogenous; G protein-coupled receptor; GPCR; receptor.
XX
OS Homo sapiens.
XX
XX US2003105292-A1.
XX
XX 05-JUN-2003.
XX
XX 20-SEP-2002; 2002US-00251385.
XX
XX 26-JUN-1998; 98US-0090783P.
XX
XX 07-AUG-1998; 98US-0095677P.
```


PR 13-OCT-1998; 98US-00170496.
XX (LIAN/) LIAN C W.
PA (BEHA/) BEHAN D P.
PA (CHAL/) CHALMERS D T.
XX
XX Liaw CW, Behan DP, Chalmers DT;
XX
XX WPI; 2003-801247/75.
DR N-PSDB; ADH14255.
XX
XX
XX New constitutively active, non-endogenous version of an endogenous human
PT G protein-coupled receptor for the identification of therapeutic
PT compounds, such as agonists.
XX
XX Example 1; SEQ ID NO 264; 227pp; English.
XX
XX The invention relates to a constitutively active, non-endogenous version
CC of an endogenous human G protein-coupled receptor (GPCR). The GPCR is
CC used for screening therapeutic compounds as inverse agonists, agonists or
CC partial agonists. The GPCR can be also used to elucidate and
CC understand the roles of GPCRs in normal and diseased humans. The GPCR
CC need not be purified and isolated to be used to screen for therapeutic
CC compounds. The utility of the GPCR as a research tool is enhanced because
CC the role of a particular receptor can be understood before the endogenous
CC ligand is identified. The present sequence is used in the exemplification
CC of the present invention.
XX
XX Sequence 907 AA;
SQ
Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLGVLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSRL 60
DB 1 MDTSLGVLSPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSRL 60
QY 61 PNLNSVFTSYLDLNMNLSQLLPNLPFLRFEELRLAGNALTYIPKGAFTGLYSKVL 120
DB 61 PNLNSVFTSYLDLNMNLSQLLPNLPFLRFEELRLAGNALTYIPKGAFTGLYSKVL 120
QY 121 LQNNQLRHVPTEALQNLRLDANHSVPPSCFSGLSRLHRLMDNALTEIPVQ 180
DB 121 LQNNQLRHVPTEALQNLRLDANHSVPPSCFSGLSRLHRLMDNALTEIPVQ 180
QY 181 AFRSLAQAMTLANKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFDGLHSL 240
DB 181 AFRSLAQAMTLANKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFDGLHSL 240
QY 241 LNNYNDLDEPPTAIRTLNLSKELGPHSNIRSIPEKAFVGNPSLIITHFYDNP 300
DB 241 LNNYNDLDEPPTAIRTLNLSKELGPHSNIRSIPEKAFVGNPSLIITHFYDNP 300
QY 301 FOHLPELRTLNGASQITEFPDLTGCTANLESITLGTGAQISSLPQVNCNQLPVL 360
DB 301 FOHLPELRTLNGASQITEFPDLTGCTANLESITLGTGAQISSLPQVNCNQLPVL 360
QY 361 YNLLEDLPSFVSCOKLOKIDIRHNEIYEIKVDTFOQLLSRLSLNLANWKIAI 420
DB 361 YNLLEDLPSFVSCOKLOKIDIRHNEIYEIKVDTFOQLLSRLSLNLANWKIAI 420
QY 421 LPSLTKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELVMPYAYOCC 480
DB 421 LPSLTKLDLSSNLSFPITGLHGLTHLKTGNHALQSLISSENPELVMPYAYOCC 480
QY 481 AFGVCENAYKISQNMKNDSMDLHKDQAGMFAQDERDLEDFLLDFEEDLKALHSV 540
DB 481 AFGVCENAYKISQNMKNDSMDLHKDQAGMFAQDERDLEDFLLDFEEDLKALHSV 540
QY 541 CSPSPGPFKPCBHLDDGMLIRIGVWTIAVLTALTCNALVTSTVFRSPLIYSP 600
DB 541 CSPSPGPFKPCBHLDDGMLIRIGVWTIAVLTALTCNALVTSTVFRSPLIYSP 600

QY 601 AVNMLTGVSASVLAGVDFTFGSPARHGAWENGCHVIGFSLTFASSSVFLTLAAL 660
DB 601 AVNMLTGVSASVLAGVDFTFGSPARHGAWENGCHVIGFSLTFASSSVFLTLAAL 660
QY 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLALTAAVPLLGSGSKYGASPLCLPLPFG 720
DB 661 ERGFSVKYSKAPETKAPFSSSLKVIILLCALLALTAAVPLLGSGSKYGASPLCLPLPFG 720
QY 721 STMGYVVALIILNSLCFLMNTIATYTKLYCNLDKGLDENIWDGSMVKHIALLLFTNCILNC 780
DB 721 STMGYVVALIILNSLCFLMNTIATYTKLYCNLDKGLDENIWDGSMVKHIALLLFTNCILNC 780
QY 781 PVAFISFSSLINLTISPVEIKFILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTV 840
DB 781 PVAFISFSSLINLTISPVEIKFILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTV 840
QY 841 WTRSKHPSLMSINSDVVKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVTESCHLSS 900
DB 841 WTRSKHPSLMSINSDVVKQSCDSTQALVTFTSSITVDLPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPECL 907
DB 901 VAFVPECL 907
RESULT 10
ADN40013
XX ADN40013 standard; protein; 907 AA.
XX AC ADN40013;
XX DT 17-JUN-2004 (first entry)
XX DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C383.
XX DE Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularization syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnerary; gene therapy; vaccine.
XX OS Homo sapiens.
XX PN WO2003042661-A2.
XX PD 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
XX PR 29-NOV-2001; 2001US-0334393P.
XX PR 03-DEC-2001; 2001US-0335394P.
XX PR 14-DEC-2001; 2001US-0340376P.
XX PR 08-JAN-2002; 2002US-0347211P.
XX PR 10-JAN-2002; 2002US-0347349P.
XX PR 08-FEB-2002; 2002US-035250P.
XX PR 13-FEB-2002; 2002US-0356714P.
XX PR 20-FEB-2002; 2002US-0359077P.
XX PR 29-MAR-2002; 2002US-036809P.
XX PR 04-APR-2002; 2002US-0370110P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PR 05-JUN-2002; 2002US-0386614P.
XX PR 16-JUL-2002; 2002US-0396839P.
XX PR 22-JUL-2002; 2002US-039775P.
XX PR 22-JUL-2002; 2002US-0397845P.
XX PR 09-SEP-2002; 2002US-0409450P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
XX

PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;
XX WPI; 2003-468649/44.
XX N-PSDB; ADN39796.
PT Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO C383; 1385pp; English.
XX
CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 907 AA;

Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSLGSEL 60
Db 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSLGSEL 60
Qy 61 PSNLSVFTSYLDLNMNNSQLPNPLPSLRFLEELRAGNALTYPKGAFTGLSKVLK 120
Db 61 PSNLSVFTSYLDLNMNNSQLPNPLPSLRFLEELRAGNALTYPKGAFTGLSKVLK 120
Qy 121 LQNNQLRHVPTEALQNLRSLSRLDANHI SYVPPSCFSLHSLRHLWLDNALTEIPVQ 180
Db 121 LQNNQLRHVPTEALQNLRSLSRLDANHI SYVPPSCFSLHSLRHLWLDNALTEIPVQ 180
Qy 181 AFRSLAQMTLALNKHITHIPDYAFGNLSLVLLHNNRTHSLGKKCFDGLHSLFTLD 240
Db 181 AFRSLAQMTLALNKHITHIPDYAFGNLSLVLLHNNRTHSLGKKCFDGLHSLFTLD 240
Qy 241 LNYNLDLFPPTAIRTLNLSKELGPHSNNSIRIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
Db 241 LNYNLDLFPPTAIRTLNLSKELGPHSNNSIRIPEKAFVGNPSLTIHFYDNP IQVGRSA 300
Qy 301 FOHLPELRTLINGASQITEPPDLTGTANLESRLTGAQISLSPQTCVQNLPNQVLDLS 360
Db 301 FOHLPELRTLINGASQITEPPDLTGTANLESRLTGAQISLSPQTCVQNLPNQVLDLS 360
Qy 361 YNLLDLPSFVSCQKQKIDLRHNEIYEIKVDFTFOQLLSLRLNLANWKIAIHPNAPST 420
Db 361 YNLLDLPSFVSCQKQKIDLRHNEIYEIKVDFTFOQLLSLRLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPVAQCC 480
Db 421 LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLISSENPELKVIEPVAQCC 480
Qy 481 AFGVCENAYKISQNWKNKGNSSMDLHKDKAGMFOQDERDLEDLDFEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGNSSMDLHKDKAGMFOQDERDLEDLDFEDLKALHSVQ 540
Qy 541 CSPSPGPKPCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600

Db 541 CSPSPGPKPCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGVSASVLAGVDAFTFGSPARHGAWWENGVGCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSASVLAGVDAFTFGSPARHGAWWENGVGCHVIGFLSIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSKAFETKAPFSSIKVILLCALIALTMAAVPLLGSSKYGASPLCLPLPGCEP 720
Db 661 ERGFSVKYSKAFETKAPFSSIKVILLCALIALTMAAVPLLGSSKYGASPLCLPLPGCEP 720
Qy 721 STMGYVVALIILNLSLCFLMTIATYKLYCNLDKGDLENIDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVVALIILNLSLCFLMTIATYKLYCNLDKGDLENIDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFSLFSSLLNLTIFISPEVVKFILLVVVPLPACINPLLYLFLNPHPKEDIVSRKQTYV 840
Db 781 PVAFSLFSSLLNLTIFISPEVVKFILLVVVPLPACINPLLYLFLNPHPKEDIVSRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVTTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDDVEKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVTTESCHLSS 900
Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907
RESULT 11
ADN39531
ID ADN39531 standard; protein; 907 AA.
XX
AC ADN39531;
XX
DT 17-JUN-2004 (first entry)
XX
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A131.
KW Human; differential expression; cancer; angiogenic disorder;
KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
KW inflammatory disease; autoimmune disease;
KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulneryary; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN WO2003042661-A2.
XX
PD 22-MAY-2003.
XX
PF 13-NOV-2002; 2002WO-US036810.
XX
PR 13-NOV-2001; 2001US-0350666P.
PR 21-NOV-2001; 2001US-0332464P.
PR 29-NOV-2001; 2001US-0334393P.
PR 03-DEC-2001; 2001US-0335394P.
PR 14-DEC-2001; 2001US-0340376P.
PR 08-JAN-2002; 2002US-0347211P.
PR 10-JAN-2002; 2002US-0347349P.
PR 08-FEB-2002; 2002US-035250P.
PR 13-FEB-2002; 2002US-0356714P.
PR 20-FEB-2002; 2002US-0359077P.
PR 29-MAR-2002; 2002US-0368809P.
PR 04-APR-2002; 2002US-0370110P.
PR 12-APR-2002; 2002US-0372246P.
PR 05-JUN-2002; 2002US-0386614P.
PR 16-JUL-2002; 2002US-0396839P.
PR 22-JUL-2002; 2002US-039775P.
PR 22-JUL-2002; 2002US-0397845P.
PR 09-SEP-2002; 2002US-0409450P.
XX

PA	(EOSB-) EOS BIOTECHNOLOGY INC.	
XX		
PI	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;	
PI	Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;	
XX		
DR	WPI: 2003-468649/44.	
DR	N-PSDB; ADN39530.	
XX		
PT	Determining the presence or absence of a pathological cell in a patient,	
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting	
PT	a nucleic acid in a biological sample.	
XX		
PS	Claim 12; SEQ ID NO A131; 1385pp; English.	
XX		
CC	The invention relates to nucleic acids and proteins (ADN38683-ADN40064)	
CC	whose expression is upregulated or downregulated in specific cancers or	
CC	other diseases such as angiogenic or fibrotic disorders, and to methods	
CC	of determining the presence or absence of a pathological cell in a	
CC	patient by detecting a nucleic acid at least 80% identical to those of	
CC	the invention or by detecting a polypeptide of the invention. The	
CC	invention also relates to expression vectors and host cells comprising a	
CC	nucleic acid of the invention; antibodies which specifically bind a	
CC	polypeptide of the invention; use of such antibodies for drug targeting;	
CC	and methods of screening for modulators of activity or expression of the	
CC	polypeptides and nucleic acids. The nucleic acids, polypeptides,	
CC	antibodies and methods are useful for diagnosing, prognosing and treating	
CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,	
CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal	
CC	neovascularisation syndromes, scarring and uterine fibroids. They may	
CC	also be useful in wound healing and in contraception. The present	
CC	sequence represents a polypeptide of the invention.	
XX		
SQ	Sequence 907 AA;	
	Query Match	
	Best Local Similarity 100.0%; Score 907; DB 7; Length 907;	
	Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGLSEL 60	
Db	1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGLSEL 60	
Qy	61 PSNLSVFTSYLDLSNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLSLKVLV 120	
Db	61 PSNLSVFTSYLDLSNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLSLKVLV 120	
Qy	121 LQNNQLRHVPTEALQNLRSQSLRDANHSVVPSPCFSGHLSLRLHLDNALTETIPVQ 180	
Db	121 LQNNQLRHVPTEALQNLRSQSLRDANHSVVPSPCFSGHLSLRLHLDNALTETIPVQ 180	
Qy	181 AFRSLSLAQMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFDGLHSLETLD 240	
Db	181 AFRSLSLAQMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFDGLHSLETLD 240	
Qy	241 LNNNLDEPPTAIRTLNKLKELGFSNNIRSIPEKAFVGNPSLTIHFYDNPIDQVGRSA 300	
Db	241 LNNNLDEPPTAIRTLNKLKELGFSNNIRSIPEKAFVGNPSLTIHFYDNPIDQVGRSA 300	
Qy	301 FOHLPELRTLINGASQITTEPDLTGTLNLSLTGTGAQISLPTQVNCNLPNQLVLDLS 360	
Db	301 FOHLPELRTLINGASQITTEPDLTGTLNLSLTGTGAQISLPTQVNCNLPNQLVLDLS 360	
Qy	361 YNLEDLPSFVSCQKQKIDLRHNEIYKVDVTFQQLSLSLSLNLANWKIATIHNPAST 420	
Db	361 YNLEDLPSFVSCQKQKIDLRHNEIYKVDVTFQQLSLSLSLNLANWKIATIHNPAST 420	
Qy	421 LPSLTKLDSNLSLSPFITGLHGLTHLKTGNHALQSLISSENPELKIEMPVAYQCC 480	
Db	421 LPSLTKLDSNLSLSPFITGLHGLTHLKTGNHALQSLISSENPELKIEMPVAYQCC 480	
Qy	481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOAQRDLFDLDFEEDLKALHSVQ 540	
Db	481 AFGVCENAYKISNOWNKGDNSMDLHKDKAGMFOAQRDLFDLDFEEDLKALHSVQ 540	

Qy	541 CSPSPGPPKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600	
Db	541 CSPSPGPPKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIA 600	
Qy	601 AVNMLTGVSAAVLAVAGVDAFTFGSFARHGAWWENGCHVIGFLSTFASSESVFLTLAAL 660	
Db	601 AVNMLTGVSAAVLAVAGVDAFTFGSFARHGAWWENGCHVIGFLSTFASSESVFLTLAAL 660	
Qy	661 ERGSVKYSKAFETKAPFSSLKVIILLCALLATMAAVPLLGSGSKYGASPLCLPFGGP 720	
Db	661 ERGSVKYSKAFETKAPFSSLKVIILLCALLATMAAVPLLGSGSKYGASPLCLPFGGP 720	
Qy	721 STMGVMAVLIILNSLCFLMMTIATYTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780	
Db	721 STMGVMAVLIILNSLCFLMMTIATYTKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC 780	
Qy	781 PVAFLSFSSLLNLNLFISPEVKIFILLVVPVLPACLNPLLYILFNPHFKEDLVSLRKQTV 840	
Db	781 PVAFLSFSSLLNLNLFISPEVKIFILLVVPVLPACLNPLLYILFNPHFKEDLVSLRKQTV 840	
Qy	841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVTESCHLSS 900	
Db	841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITYDLPPSSVPSPAYPVTESCHLSS 900	
Qy	901 VAFVPCPL 907	
Db	901 VAFVPCPL 907	
RESULT 12		
ADN39628		
ID	ADN39628 standard; protein; 907 AA.	
XX		
AC	ADN39628;	
XX		
DT	17-JUN-2004 (first entry)	
XX		
DE	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:A228.	
KW	Human; differential expression; cancer; angiogenic disorder;	
KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;	
KW	inflammatory disease; autoimmune disease;	
KW	retinal neovascularisation syndrome; scarring; uterine fibroid;	
KW	detection; diagnosis; prognosis; drug screening; drug targeting;	
KW	wound healing; contraception; cytostatic; cardiant; immunomodulatory;	
KW	vulnery; gene therapy; vaccine.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2003042661-A2.	
XX		
PD	22-MAY-2003.	
XX		
PF	13-NOV-2002; 2002WO-US036810.	
XX		
PR	13-NOV-2001; 2001US-0350666P.	
PR	21-NOV-2001; 2001US-0332464P.	
PR	29-NOV-2001; 2001US-0334393P.	
PR	03-DEC-2001; 2001US-0335394P.	
PR	14-DEC-2001; 2001US-0340376P.	
PR	08-JAN-2002; 2002US-0347211P.	
PR	10-JAN-2002; 2002US-0347349P.	
PR	08-FEB-2002; 2002US-035250P.	
PR	13-FEB-2002; 2002US-0356714P.	
PR	20-FEB-2002; 2002US-0359077P.	
PR	29-MAR-2002; 2002US-0368809P.	
PR	04-APR-2002; 2002US-0370110P.	
PR	12-APR-2002; 2002US-0372246P.	
PR	05-JUN-2002; 2002US-0386614P.	
PR	16-JUL-2002; 2002US-0396839P.	
PR	22-JUL-2002; 2002US-0397775P.	
PR	22-JUL-2002; 2002US-0397845P.	

PR 09-SEP-2002; 2002US-0409450P.
XX (ROSB-) EOS BIOTECHNOLOGY INC.
XX Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
XX WPI; 2003-468649/44.
DR N-PSDB; ADN39627.
XX
XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.
XX
PS Claim 12; SEQ ID NO A228; 1385pp; English.
XX
XX The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.
XX
SQ Sequence 907 AA;
Query Match 100.0%; Score 907; DB 7; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
Db 1 MDTSLRGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
QY 61 PSNLVFTSYDLDSNNISQLLPNLPFLRFLRAGNALTYPKGAFTGLYSLKVL 120
Db 61 PSNLVFTSYDLDSNNISQLLPNLPFLRFLRAGNALTYPKGAFTGLYSLKVL 120
QY 121 LQNNQLRHVPTALQNLRSQSLRDANHSVPPSCFSGLSLRLHMLDDNALTEIPVQ 180
Db 121 LQNNQLRHVPTALQNLRSQSLRDANHSVPPSCFSGLSLRLHMLDDNALTEIPVQ 180
QY 181 AFRSLALQAMTALANKIHIPDYAFGNLSLVVLHNNRIHSLGKCFDGLHSLTLD 240
Db 181 AFRSLALQAMTALANKIHIPDYAFGNLSLVVLHNNRIHSLGKCFDGLHSLTLD 240
QY 241 LNNYNDLDEPPTAIRLSNKLGFHNNIRSIPEKAFVGNPNSLIITHFYDNPQVGRSA 300
Db 241 LNNYNDLDEPPTAIRLSNKLGFHNNIRSIPEKAFVGNPNSLIITHFYDNPQVGRSA 300
QY 301 FOHLPELRTLUNGASQITERPDLTGANTLESILTGAQISSLPQVCNQLPNLQVLDLS 360
Db 301 FOHLPELRTLUNGASQITERPDLTGANTLESILTGAQISSLPQVCNQLPNLQVLDLS 360
QY 361 YNLLDLPSFSVCQKLOKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAST 420
Db 361 YNLLDLPSFSVCQKLOKIDLRHNEIYEIKVDTFQQLSLRSLNLANWKIAIHPNAST 420
QY 421 LPSLKLDSLNNLSFPITGLHGLTHLKTGNHALQSLISSENPELKIEMPIAYOCC 480
Db 421 LPSLKLDSLNNLSFPITGLHGLTHLKTGNHALQSLISSENPELKIEMPIAYOCC 480
QY 481 AFGVCENAYKISNQWNGKNSMDLHKDAGMFQAQDERDLEDFLLDFEEDLKALHSVQ 540

Db 481 AFGVCENAYKISNQWNGKNSMDLHKDAGMFQAQDERDLEDFLLDFEEDLKALHSVQ 540
QY 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPFKPCHEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
QY 601 AVNMLTGVSAAVLGADAFPGSFARHCAMWENGVCCHVIGFLSTIFASESSVELLTLAAL 660
Db 601 AVNMLTGVSAAVLGADAFPGSFARHCAMWENGVCCHVIGFLSTIFASESSVELLTLAAL 660
QY 661 ERGFSVKYSKAFETKAPFSSLKVIILLCALIALTAAVPLLGSKYKASPLCLPFGCEP 720
Db 661 ERGFSVKYSKAFETKAPFSSLKVIILLCALIALTAAVPLLGSKYKASPLCLPFGCEP 720
QY 721 STMGYVALIILNLSLCFLMMTIATKYCNLDKGDLENINWDCSMVKHIALLLFNLCILNC 780
Db 721 STMGYVALIILNLSLCFLMMTIATKYCNLDKGDLENINWDCSMVKHIALLLFNLCILNC 780
QY 781 PVAFLSFSSILNLTPIISPEVTKFLLVVVPLPACINPLLYILFNPHFKEDLVSLRKQIV 840
Db 781 PVAFLSFSSILNLTPIISPEVTKFLLVVVPLPACINPLLYILFNPHFKEDLVSLRKQIV 840
QY 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPPSSVPSPAYPVTESCHLSS 900
QY 901 VAFVPC 907
Db 901 VAFVPC 907
RESULT 13
ADO29408
ID ADO29408 standard; protein; 907 AA.
XX AC ADO29408;
XX DT 29-JUL-2004 (first entry)
XX DE Human GPCR GPR49, SEQ ID NO:510.
KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antianemic;
KW cytostatic; antinflammatory; vasotropic; antiangiogenic; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW viricide; hepatotropic; antibacterial; antianaemic; antiseborrhoeic;
KW dermatological; antitumor; antithyroid; antiallergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW receptor.
XX Homo sapiens.
XX WO2004040000-A2.
PD 13-MAY-2004.
XX 09-SEP-2003; 2003WO-US028226.
XX 09-SEP-2002; 2002US-0409303P.
PR 09-APR-2003; 2003US-0461329P.
XX (PRIM-) PRIMAL INC.
PI Gaitanaris GA, Bergmann JF, Gragerov A, Hohmann J, Li P;
PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;

XX WPI; 2004-390329/36.
DR N-PSDB; ADO29923.
XX
XX Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX
XX Claim 151; SEQ ID NO 510; 542pp; English.
XX
XX The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridize to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); immune disorders (e.g., autoimmune disorders or
CC anaemia or leukaemia); AIDS; bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
CC invention. Note: The full sequence data for this patent did not form part
CC of the printed specification; those sequences not shown were obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 907 AA;

Query Match 100.0%; Score 907; DB 8; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB |||||
QY 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL 60
DB |||||
QY 61 PNLNLSVFTSYLDLNMNLSQLLPDLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLUM 120
DB |||||
QY 61 PNLNLSVFTSYLDLNMNLSQLLPDLPSLRFLEELRAGNALTYIPKGAFTGLYSKVLUM 120
DB |||||
QY 121 LQNNQLRHVPTALQNLRLQSLRLDANHSVVPSPCSGLSLRHLMDLDDNALTEIPVQ 180
DB |||||
QY 121 LQNNQLRHVPTALQNLRLQSLRLDANHSVVPSPCSGLSLRHLMDLDDNALTEIPVQ 180
DB |||||
QY 181 AFRSLSAQAMTLALNKIHHIPDYAFGNLSLVLHLHNNRIHSLGKKFCGLHSLETLD 240
DB |||||
QY 181 AFRSLSAQAMTLALNKIHHIPDYAFGNLSLVLHLHNNRIHSLGKKFCGLHSLETLD 240
DB |||||
QY 241 LNNYNDLDEPTAIRTLNLSKELGFHSNNIRSIPEKAFVGNPNSLIITHFYDNPDIQVGRSA 300
DB |||||
QY 241 LNNYNDLDEPTAIRTLNLSKELGFHSNNIRSIPEKAFVGNPNSLIITHFYDNPDIQVGRSA 300
DB |||||
QY 301 FOHLPELRPLTLNGASQITFEFPLDTCTANLESILTLTGAQISLPQVCNQLPNLQVLDLS 360
DB |||||
QY 301 FOHLPELRPLTLNGASQITFEFPLDTCTANLESILTLTGAQISLPQVCNQLPNLQVLDLS 360
DB |||||

QY 361 YNLEDLPSFSVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHPNAST 420
DB |||||
QY 361 YNLEDLPSFSVCQKLOKIDLRHNEIYEIKVDTFOQLLSRLSLNLAWNKIAIHPNAST 420
DB |||||
QY 421 LPSLKLKLDLSSNLSFPITGLHGLTHLKLCTGNHALQSLISSENPPKELKVIEMPIAYQCC 480
DB |||||
QY 421 LPSLKLKLDLSSNLSFPITGLHGLTHLKLCTGNHALQSLISSENPPKELKVIEMPIAYQCC 480
DB |||||
QY 481 AFGVCENAYKISQNWKNKGDNSMDDLHKKQAGMFOAQDQDERDLEDFLDDFEEDKALHSVQ 540
DB |||||
QY 481 AFGVCENAYKISQNWKNKGDNSMDDLHKKQAGMFOAQDQDERDLEDFLDDFEEDKALHSVQ 540
DB |||||
QY 541 CSPSPGPKPCCHLLDGLHILIRIGVMTIAVLAITCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB |||||
QY 541 CSPSPGPKPCCHLLDGLHILIRIGVMTIAVLAITCNALVTSTVFRSPLYISPIKLLIGVIA 600
DB |||||
QY 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660
DB |||||
QY 601 AVNMLTGVSASVAVLAGVDAFTFGSPARHGAWNGVCHVIGFLSIFASESSVFLTLAAL 660
DB |||||
QY 661 ERGFSVKYSAKPETKAPFSSSLKVIILLCALLALTMAAVPLLGSKYKYGASPLCLPLPFGEP 720
DB |||||
QY 661 ERGFSVKYSAKPETKAPFSSSLKVIILLCALLALTMAAVPLLGSKYKYGASPLCLPLPFGEP 720
DB |||||
QY 721 STMGYVVALIILNSLCFLMWTIATYKLYCNLDKGLDENIWDSCMWKHIALLFTNCILNC 780
DB |||||
QY 721 STMGYVVALIILNSLCFLMWTIATYKLYCNLDKGLDENIWDSCMWKHIALLFTNCILNC 780
DB |||||
QY 781 PVAFSLSPSLINLTFTISPEVIKFIILVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTVV 840
DB |||||
QY 781 PVAFSLSPSLINLTFTISPEVIKFIILVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTVV 840
DB |||||
QY 841 WTRSKHPSLMSINSDDVKQSCDSTQALVFTTSSSITVDLPSPSSVPSPAYPVTESCHLSS 900
DB |||||
QY 841 WTRSKHPSLMSINSDDVKQSCDSTQALVFTTSSSITVDLPSPSSVPSPAYPVTESCHLSS 900
DB |||||
QY 901 VAFVPCCL 907
DB |||||
QY 901 VAFVPCCL 907
DB |||||

RESULT 14
ID ADQ80369
AD ADQ80369 standard; protein; 907 AA.
XX
XX AC ADQ80369;
XX
XX DT 21-OCT-2004 (first entry)
XX
XX DE G protein-coupled receptor 49 protein.
XX
XX KW cytostatic; epidermal growth factor receptor modulator; identification;
XX therapeutic response; cancer; EGFR; biomarker.
XX
XX OS Homo sapiens.
XX
XX PN WO2004063709-A2.
XX
XX PD 29-JUL-2004.
XX
XX PF 08-JAN-2004; 2004WO-US0000368.
XX
XX PR 08-JAN-2003; 2003US-0438735P.
XX
XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX PI Amler LC, Januario T;
XX
XX DR WPI; 2004-544114/52.
XX
XX DR N-PSDB; ADQ80249.
XX
XX PT Identifying a mammal that will respond therapeutically to a method of
PT treating cancer comprises comparing the level of a biomarker in a mammal

PT before and after exposure to an epidermal growth factor receptor (EGFR)
PT modulator.

PS Disclosure; SEQ ID NO 141; 520pp; English.

XX The invention relates to a method of identifying a mammal that will
CC respond therapeutically to a method of treating cancer by administering
CC an epidermal growth factor receptor (EGFR) modulator by comparing the
CC level of a biomarker in a mammal before and after exposure to an EGFR
CC modulator. The method comprises: (a) measuring, in the mammal, the level
CC of at least one biomarker identified in the specification; (b) exposing
CC the mammal to the EGFR modulator; and (c) measuring in the mammal the
CC level of the biomarker, where a difference in the level in step (c)
CC compared to step (a) indicates that the mammal will respond
CC therapeutically to the method of treating cancer. The method and
CC biomarkers are useful for identifying a mammal that will respond
CC therapeutically to a method of treating cancer by administering an
CC epidermal growth factor receptor (EGFR) modulator. This sequence
CC corresponds to one of the biomarkers whose levels of expression is
CC measured in the method of the invention.

XX Sequence 907 AA;

Query Match	100.0%;	Score 907;	DB 8;	Length 907;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 907;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDTSRLGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSDIGLSEL	60	
Db	1	MDTSRLGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCPDGRMLLRVDCSDIGLSEL	60	
QY	61	PSNLSVFTSYDLDSNNISQLLPNLPSPILFLEELRLAGNALTYIPKGAFTGLYSKVLV	120	
Db	61	PSNLSVFTSYDLDSNNISQLLPNLPSPILFLEELRLAGNALTYIPKGAFTGLYSKVLV	120	
QY	121	LQNNQLRHVPTEALQNLASLRLDANHSIVPPSCFSGLSLHRLMLDNLALTEIPVQ	180	
Db	121	LQNNQLRHVPTEALQNLASLRLDANHSIVPPSCFSGLSLHRLMLDNLALTEIPVQ	180	
QY	181	APRSLSALQAMTALANKIHHIPDYAFGNLSLVVLHNNRIHSLGKCKPGLHSLFETLD	240	
Db	181	APRSLSALQAMTALANKIHHIPDYAFGNLSLVVLHNNRIHSLGKCKPGLHSLFETLD	240	
QY	241	LYNNNLDEFPFAIRTLNKLGFHNSNIRSTRPEKAFVGNPSLTIHFYDNPFIQVGRSA	300	
Db	241	LYNNNLDEFPFAIRTLNKLGFHNSNIRSTRPEKAFVGNPSLTIHFYDNPFIQVGRSA	300	
QY	301	FQHLPELRTLTLNGASQITEFPDLTGTLNLSLTLTGAQISLSPQTVCNQPNLQVLDLS	360	
Db	301	FQHLPELRTLTLNGASQITEFPDLTGTLNLSLTLTGAQISLSPQTVCNQPNLQVLDLS	360	
QY	361	YNLLEDLPSFVCQKLOKIDLRHNEIYEIKVDTPQOOLSLSLNLANWKIAIHHNAPST	420	
Db	361	YNLLEDLPSFVCQKLOKIDLRHNEIYEIKVDTPQOOLSLSLNLANWKIAIHHNAPST	420	
QY	421	LPSLTKLDLSSNLSSEFPITGLHGTTHLKTGNHALQSLISENPELKIEMPVAYOCC	480	
Db	421	LPSLTKLDLSSNLSSEFPITGLHGTTHLKTGNHALQSLISENPELKIEMPVAYOCC	480	
QY	481	AFGVENAYKISQNNWKGDNSSMDLHKDQAGMFOAQDERDLEDFLLDFEEDLKALHSVQ	540	
Db	481	AFGVENAYKISQNNWKGDNSSMDLHKDQAGMFOAQDERDLEDFLLDFEEDLKALHSVQ	540	
QY	541	CSPSPGPKPCBHLDDGMLIRIGVWTVIAVLTALTCNALVTSTVFRSPLYSIPKLLIGVIA	600	
Db	541	CSPSPGPKPCBHLDDGMLIRIGVWTVIAVLTALTCNALVTSTVFRSPLYSIPKLLIGVIA	600	
QY	601	AVNMLTGVSASVAGVDAFTGSGFARHGAWENGVCVIGPLSTFASSESVFLTLTAAAL	660	
Db	601	AVNMLTGVSASVAGVDAFTGSGFARHGAWENGVCVIGPLSTFASSESVFLTLTAAAL	660	
QY	661	ERGSFVKYSAKFETKAPFSSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP	720	

Db	661	ERGSFVKYSAKFETKAPFSSSLKVIILLCALLALTMAAVPLLGSGSKYGASPLCLPLPFGEP	720
QY	721	STMGYVALIILNLSICFLMMTIATKLYCNLDKGDLENINWDCSMVKHTALLFTNCILNC	780
Db	721	STMGYVALIILNLSICFLMMTIATKLYCNLDKGDLENINWDCSMVKHTALLFTNCILNC	780
QY	781	PVAFLSFSSLINLTFISPEVIKFTLLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTYV	840
Db	781	PVAFLSFSSLINLTFISPEVIKFTLLVVVPLPACINPLLYILFNPHFKEDLVSLRKQTYV	840
QY	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSPSPVPAYPVWTSCHLSS	900
Db	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTFTSSITVDLPSPSPVPAYPVWTSCHLSS	900
QY	901	VAFVPCPL 907	
Db	901	VAFVPCPL 907	

RESULT 15

ADR67868
ID ADR67868 standard; protein; 907 AA.

XX ADR67868;
AC ADR67868;
DT 18-NOV-2004 (first entry)
XX Human HG38 protein.
DE human; G protein-coupled receptor; GPCR; HG38; colon; lung; cancer.
KW Homo sapiens.
OS Homo sapiens.
PN WO2004074436-A2.
XX 02-SEP-2004.
XX 11-FEB-2004; 2004WO-US004060.
PF 19-FEB-2003; 2003US-0448959P.
PR (INCY-) INCYTE CORP.

XX Lasek AW;
XX WPI; 2004-652946/63.
DR N-PSDB; ADR67869.

Detecting colon or lung cancer, by determining amount of protein in sample, comparing amount of protein to standard, and differential expression of protein in sample indicates colon or lung cancer.

Claim 4; SEQ ID NO 1; 79pp; English.

This sequence represents the human G protein-coupled receptor (GPCR) known as HG38. The HG38 protein and corresponding nucleic acid, may be used in the method of the invention for detecting colon or lung cancer. The first method involves performing an assay to determine the amount of HG38 in a sample of colon or lung tissue, and comparing the amount of protein to standard, thus detecting expression of protein in sample, where differential expression of protein in sample when compared with the standard is diagnostic of colon or lung cancer. The second method involves hybridizing a composition comprising the HG38 coding sequence, or its complement, and a labelling moiety, to nucleic acids of a sample of colon or lung tissue under conditions to form at least one hybridization complex, detecting hybridization complex formation, and comparing complex formation to a standard, where the comparison reflects differential expression of the polynucleotide in the sample relative to the standard and is diagnostic of a colon or lung cancer. This first method enables earlier diagnosis before the patient is symptomatic. Anti-HG38 antibodies are useful for treating colon or lung cancer.

Sequence 907 AA;

Query Match		100.0%;	Score 907;	DB 8;	Length 907;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 907;	Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;
Qy	1	MDTSRLGVLLSLPVLLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLRVDCSDGLSEL	60		
Db	1	MDTSRLGVLLSLPVLLQLATGSSPRSGVLLRGCTHCHCEPDGRMLLRVDCSDGLSEL	60		
Qy	61	PSNLSVFTSYLDLSMNNISQLPNLPSLRPLEELRLAGNALTYIPKGAFTGLYSLKVLV	120		
Db	61	PSNLSVFTSYLDLSMNNISQLPNLPSLRPLEELRLAGNALTYIPKGAFTGLYSLKVLV	120		
Qy	121	LQNNQLRHVPTEALQNLRLSLQSLRLDANHISYVPPSCFSGLSLRLMLDDNALTEIPVQ	180		
Db	121	LQNNQLRHVPTEALQNLRLSLQSLRLDANHISYVPPSCFSGLSLRLMLDDNALTEIPVQ	180		
Qy	181	AFRSLSALQAMTLALNKIHHPDYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLTLD	240		
Db	181	AFRSLSALQAMTLALNKIHHPDYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLTLD	240		
Qy	241	LNYNLDRFPFAIRTLNKLGFHSHNNIRSIPEKAFVGNPSLIITHFYDNP IQFVGRSA	300		
Db	241	LNYNLDRFPFAIRTLNKLGFHSHNNIRSIPEKAFVGNPSLIITHFYDNP IQFVGRSA	300		
Qy	301	FQHLPELRTLTLNGASQITEPPDLTG TANLES LTLTGAQISSLPQTVCNQPNLQVLDLS	360		
Db	301	FQHLPELRTLTLNGASQITEPPDLTG TANLES LTLTGAQISSLPQTVCNQPNLQVLDLS	360		
Qy	361	YNLEDLPSFVCQKQKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIAIHHENAFST	420		
Db	361	YNLEDLPSFVCQKQKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIAIHHENAFST	420		
Qy	421	LPSLIKLDSLNNLSSFPITGLHGTJHLKLTGNHALQSLISSENPELKVTEMPIYAYOCC	480		
Db	421	LPSLIKLDSLNNLSSFPITGLHGTJHLKLTGNHALQSLISSENPELKVTEMPIYAYOCC	480		
Qy	481	AFGCENAYKISNQWKNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ	540		
Db	481	AFGCENAYKISNQWKNKGNSSMDLHKDKAGMFOAQDERDLEDFLLDFEEDLKALHSVQ	540		
Qy	541	CSPSPGPKPCHEHLLDGLHRIIGVWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA	600		
Db	541	CSPSPGPKPCHEHLLDGLHRIIGVWTIAVLTALTCNALVTSTVFRSPLYISPIKLLIGVIA	600		
Qy	601	AVNMLTGVSVAVLGADAFSGSFARHGAHWENGCHVIGFLSTFASSESVFLTLAAL	660		
Db	601	AVNMLTGVSVAVLGADAFSGSFARHGAHWENGCHVIGFLSTFASSESVFLTLAAL	660		
Qy	661	ERGSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGE	720		
Db	661	ERGSVKYSKAPETKAPFSSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGE	720		
Qy	721	STMGYVVALIILNLSLCFLLMTIATYKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780		
Db	721	STMGYVVALIILNLSLCFLLMTIATYKLYCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780		
Qy	781	PVAFLSFSSLNLNLTISPSEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYV	840		
Db	781	PVAFLSFSSLNLNLTISPSEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYV	840		
Qy	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSSITYDLPSSVPSPAYPVTESCHLSS	900		
Db	841	WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSSITYDLPSSVPSPAYPVTESCHLSS	900		
Qy	901	VAFVPC	907		
Db	901	VAFVPC	907		

Search completed: July 12, 2005, 07:51:17
Job time : 112 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model
Run on: July 12, 2005, 07:38:42 ; Search time 29 Seconds
(without alignments)
3009.262 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 907
Sequence: 1 MDTSLRGVLLSLPVLQLAT.....PAYPTESCHLSVAFVPC 907

Scoring table: OLIGO 60.0 , Gapext 60.0
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues
Word size : 0
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	2	JG0176
2	51	5.6	907	2	JG0193
3	9	1.0	195	2	AF2638
4	9	1.0	209	2	G97420
5	9	1.0	580	2	S06057
6	9	1.0	869	2	A71400
7	9	1.0	1978	2	S77257
8	8	0.9	162	2	F97440
9	8	0.9	178	2	A87605
10	8	0.9	179	2	S75434
11	8	0.9	238	2	AD1368
12	8	0.9	267	2	AF3511
13	8	0.9	293	2	B71686
14	8	0.9	299	2	S61248
15	8	0.9	299	2	A75591
16	8	0.9	324	2	T20679
17	8	0.9	329	2	H64143
18	8	0.9	402	2	T15490
19	8	0.9	431	2	H84392
20	8	0.9	467	2	D86583
21	8	0.9	467	2	E72040
22	8	0.9	478	2	S75572
23	8	0.9	481	2	AD3020
24	8	0.9	481	2	E98264
25	8	0.9	538	1	D84164
26	8	0.9	671	2	D84648
27	8	0.9	674	2	JC5104
28	8	0.9	678	2	C97114
29	8	0.9	729	2	F86308

30	8	0.9	732	2	S25995	hypothetical prote
31	8	0.9	851	2	S67285	NUDI protein - yea
32	8	0.9	967	2	T48210	hypothetical prote
33	8	0.9	976	2	B84659	probable receptor-
34	8	0.9	980	2	T05414	protein kinase hom
35	8	0.9	1159	1	A44280	inner layer protei
36	8	0.9	1286	2	T33476	hypothetical prote
37	8	0.9	1614	2	T29861	hypothetical prote
38	8	0.9	2025	2	T03884	hypothetical prote
39	7	0.8	94	2	E86889	hypothetical prote
40	7	0.8	98	2	A58889	NADH2 dehydrogenas
41	7	0.8	103	2	AE3419	hypothetical prote
42	7	0.8	105	4	S59325	hypothetical prote
43	7	0.8	106	2	T04804	hypothetical prote
44	7	0.8	106	2	E83435	hypothetical prote
45	7	0.8	108	2	S75540	hypothetical prote

ALIGNMENTS

RESULT 1

JB0176
orphan G protein-coupled receptor precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
C:Accession: JB0176
R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.
Biochem. Biophys. Res. Commun. 247, 266-270, 1998
A:Title: Identification and cloning of an orphan G protein-coupled receptor of the glyc
A:Reference number: JB0176; MUID:98308104; PMID:9642114
A:Accession: JB0176
A:Molecule type: mRNA
A:Residues: 1-907 <MCD>
A:Cross-references: UNIPROT-075473; GB:AF062006; NID:G3366801; PIDN:AAC28019,1; PID:G336
C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.
C:Genetics:
A:Gene: HG38
A:Map position: 12q22-23
F:1-21/Domain: signal sequence #status predicted <SIG>
F:562-583/Domain: transmembrane #status predicted <TM1>
F:594-616/Domain: transmembrane #status predicted <TM2>
F:633-660/Domain: transmembrane #status predicted <TM3>
F:681-701/Domain: transmembrane #status predicted <TM4>
F:725-744/Domain: transmembrane #status predicted <TM5>
F:768-791/Domain: transmembrane #status predicted <TM6>
F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match 100.0%; Score 907; DB 2; Length 907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 907; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MDTSLRGVLLSLPVLQLATCGSSPRSGVLRGCPETHCHCEPDGEMLRVDCSDGLSEL	60
Db	1	MDTSLRGVLLSLPVLQLATCGSSPRSGVLRGCPETHCHCEPDGEMLRVDCSDGLSEL	60
Qy	61	PSNLSVFTSYLDLSNNNISQLLPNPLPSLRLEELRAGNALTYIPKGAFTGLYSKVLV	120
Db	61	PSNLSVFTSYLDLSNNNISQLLPNPLPSLRLEELRAGNALTYIPKGAFTGLYSKVLV	120
Qy	121	LQNNQLRHVPTEALQNLRLSLQSLRDANHIISVPPSPCSGLHSRLHMLDNNALTEIPVQ	180
Db	121	LQNNQLRHVPTEALQNLRLSLQSLRDANHIISVPPSPCSGLHSRLHMLDNNALTEIPVQ	180
Qy	181	AFRSLALQAMTLAKIKIHIPDYAFGNLSLVVHLHNNRHSLSGKCKCFGLHSLETLD	240
Db	181	AFRSLALQAMTLAKIKIHIPDYAFGNLSLVVHLHNNRHSLSGKCKCFGLHSLETLD	240
Qy	241	LYNNNLEFPPTAIRTLNKLGFHSHNIRIPEKAFVGNPSLTIHFYDNPFIQVGRSA	300
Db	241	LYNNNLEFPPTAIRTLNKLGFHSHNIRIPEKAFVGNPSLTIHFYDNPFIQVGRSA	300
Qy	301	FQHLPELRTLTLNGASQITEFPDPTGTANLESRTLTAQAQISSLPQTVCNQLPNLQVLDLS	360

Db	301	FOHLELRLTLNGASQITEPDLTGTAANLSLTLTGAIQISLPQVCNQLNQLVLDLS	360
Qy	361	YNLEEDLPSFVCQKLOKIDLRHNEIYRIKVDTFQOLLSLRSLNLAWNKIAIIHPNAPST	420
Db	361	YNLEEDLPSFVCQKLOKIDLRHNEIYRIKVDTFQOLLSLRSLNLAWNKIAIIHPNAPST	420
Qy	421	LPSLIKDLSSNLLSPFITGLHGLTHLKTGNHALQSLISENPPELKVIEMPYAYOCC	480
Db	421	LPSLIKDLSSNLLSPFITGLHGLTHLKTGNHALQSLISENPPELKVIEMPYAYOCC	480
Qy	481	AFVCENAYKISQNWKNKGNSSMDLHKDAGMFOQDERDLEDLLDFEEDLKALHSVQ	540
Db	481	AFVCENAYKISQNWKNKGNSSMDLHKDAGMFOQDERDLEDLLDFEEDLKALHSVQ	540
Qy	541	CSPSPGPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA	600
Db	541	CSPSPGPKPCBHLDDGWLIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA	600
Qy	601	AVNMLTGYSVAVLGADVATFGSFARHGAWENGVCVIGFLSIFASESSVFLITLAAL	660
Db	601	AVNMLTGYSVAVLGADVATFGSFARHGAWENGVCVIGFLSIFASESSVFLITLAAL	660
Qy	661	ERGSFVKYSAKPETKAPFSSLKVIIILLCALLATMAAVPLLGGSKYGASPLCLPLPFGEF	720
Db	661	ERGSFVKYSAKPETKAPFSSLKVIIILLCALLATMAAVPLLGGSKYGASPLCLPLPFGEF	720
Qy	721	STMGYMVALIILNSLCFLMWTIATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC	780
Db	721	STMGYMVALIILNSLCFLMWTIATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC	780
Qy	781	PVAFLSFSLNLTPISEVVKFILLVVVPLPACLNPLLYLILFNPHKEDLVSLRKQTVV	840
Db	781	PVAFLSFSLNLTPISEVVKFILLVVVPLPACLNPLLYLILFNPHKEDLVSLRKQTVV	840
Qy	841	WTRSKHPSLMSINDSDVKQSCDSTQALVTFITSSITVDLPPSSVPSPAYPVPTESCHLSS	900
Db	841	WTRSKHPSLMSINDSDVKQSCDSTQALVTFITSSITVDLPPSSVPSPAYPVPTESCHLSS	900
Qy	901	VAFVPCL 907	
Db	901	VAFVPCL 907	
RESULT 2			
JG0193			
G protein-coupled receptor FXR - mouse			
C:Species: Mus musculus (house mouse)			
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000			
C:Accession: JG0193			
R:Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.			
Biochem. Biophys. Res. Commun. 254, 273-279, 1999			
A:Title: Identification of a novel seven-transmembrane receptor with homology to glycophorin A			
A:Reference number: JG0193; MUID:99121227; PMID:9920770			
A:Accession: JG0193			
A:Status: preliminary			
A:Molecule type: mRNA			
A:Residues: 1-907 <HER>			
Query Match			
Best Local Similarity 5.6%; Score 51; DB 2; Length 907;			
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	203	DYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLFTLDLNNYNNLDEFPPTAI	253
Db	203	DYAFGNLSLVVHLHNNRIHSLGKCFDGLHSLFTLDLNNYNNLDEFPPTAI	253
RESULT 3			
AF2638			
conserved hypothetical protein Atu0506 [imported] - Agrobacterium tumefaciens (strain C58)			
C:Species: Agrobacterium tumefaciens			
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004			

C:Accession: AF2638			
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.			
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan			
; Karp, P.; Romero, P.; Zhang, S.			
Science 294, 2317-2323, 2001			
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, E.W.			
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.			
A:Reference number: AB2577; MUID:21608550; PMID:11743193			
A:Accession: AF2638			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-195 <KUR>			
A:Cross-references: UNIPROT:Q8UHZ5; GB:AB008688; PIDN:AAL41524.1; PID:gi7738854; GSPDB:G			
A:Experimental source: strain C58 (dupont)			
C:Genetics:			
A:Gene: Atu0506			
A:Map position: circular chromosome			
Query Match			
Best Local Similarity 1.0%; Score 9; DB 2; Length 195;			
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	25	PRSGVLLRG 33	
Db	145	PRSGVLLRG 153	
RESULT 4			
G97420			
hypothetical protein AGR_C_894 [imported] - Agrobacterium tumefaciens (strain C58, Cerezo)			
C:Species: Agrobacterium tumefaciens			
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004			
C:Accession: G97420			
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;			
Science 294, 2323-2328, 2001			
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens			
A:Reference number: A97359; MUID:21608551; PMID:11743194			
A:Accession: G97420			
A:Status: preliminary			
A:Molecule type: DNA			
A:Residues: 1-209 <KUR>			
A:Cross-references: UNIPROT:Q8UHZ5; GB:AE007869; PIDN:AAK86320.1; PID:gi15155439; GSPDB:G			
C:Genetics:			
A:Gene: AGR_C_894			
A:Map position: circular chromosome			
Query Match			
Best Local Similarity 1.0%; Score 9; DB 2; Length 209;			
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	25	PRSGVLLRG 33	
Db	159	PRSGVLLRG 167	
RESULT 5			
S06057			
gene ND1 intron 3 protein 2 - Podospora anserina mitochondrion			
C:Species: mitochondrion Podospora anserina			
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004			
C:Accession: S06057			
R:Cummings, D.J.; Domenico, J.M.; Michel, F.			
Curr. Genet. 14, 253-264, 1988			
A:Title: DNA sequence and organization of the mitochondrial ND1 gene from Podospora anserina			
A:Reference number: S06056; MUID:89063443; PMID:3197134			
A:Accession: S06057			
A:Molecule type: DNA			
A:Residues: 1-580 <CUM>			
A:Cross-references: UNIPROT:Q02714; EMBL:X13164			
C:Genetics:			
A:Genome: mitochondrion			

A;Genetic code: SGC3
C;Keywords: mitochondrion

Query Match 1.0%; Score 9; DB 2; Length 580;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 530 BEDLKALHS 538
|||
Db 108 BEDLKALHS 116

RESULT 6
A71400
probable disease resistance protein - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cross)
A;Variety: columbia
C;Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
C;Accession: A71400
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A;Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoft, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ana
C.; Chalwatzis, N.
A;Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal
A;Reference number: A71400; MUID:98121113; PMID:9461215
A;Accession: A71400
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-869 <BEV>
A;Cross-references: UNIPROT:O23253; GB:Z97335; NID:g2244747; PID:e326857; PID:g2244748
C;Genetics:
A;Map position: 4COP9-4G3845

Query Match 1.0%; Score 9; DB 2; Length 869;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 421 LPSLIKLDL 429
|||
Db 334 LPSLIKLDL 342

RESULT 7
S77257
hypothetical protein sll1265 - Synechocystis sp. (strain PCC 6803)
C;Species: Synechocystis sp.
A;Variety: PCC 6803
C;Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C;Accession: S77257
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A;Reference number: S74322; MUID:97061201; PMID:8905231
A;Accession: S77257
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-1978 <KAN>
A;Cross-references: UNIPROT:P73551; EMBL:D90907; GB:AB001339; NID:g1652618; PIDN:BAAL759
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 1.0%; Score 9; DB 2; Length 1978;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 333 LTLTGAQIS 341
|||
Db 526 LTLTGAQIS 534

RESULT 8

F97440
hypothetical protein AGR_C_1201 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: F97440
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, A.;
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: F97440
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <KUR>
A;Cross-references: UNIPROT:Q8U5I8; GB:AE007869; PIDN:AAK86479.1; PID:g15155627; GSPDB:G
C;Genetics:
A;Gene: AGR_C_1201
A;Map position: circular chromosome

Query Match 0.9%; Score 8; DB 2; Length 162;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 690 LLALTMAA 697
|||
Db 22 LLALTMAA 29

RESULT 9

A87605
Dps family protein [imported] - Caulobacter crescentus
C;Species: Caulobacter crescentus
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C;Accession: A87605
R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A;Title: Complete Genome Sequence of Caulobacter crescentus.
A;Reference number: A87249; MUID:21173698; PMID:11259647
A;Accession: A87605
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-178 <STO>
A;Cross-references: UNIPROT:Q9A4G1; GB:AE005673; NID:g13424489; PIDN:AAK24837.1; GSPDB:G
C;Genetics:
A;Gene: CC2873

Query Match 0.9%; Score 8; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 205 AFGMLSSL 212
|||
Db 104 AFGMLSSL 111

RESULT 10

S75434
hypothetical protein c05017 - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004
C;Accession: S75434
R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S.
Mol. Microbiol. 22, 175-191, 1996
A;Title: Organizational characteristics and information content of an archaeal genome: 1
A;Reference number: S73076; MUID:97055432; PMID:8899719
A;Accession: S75434
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-179 <SEN>

A;Cross-references: UNIPROT:P96006; EMBL:Y08257; NID:gl707772; PID:c283869; PID:gl707840
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C;Genetics:
A;Start codon: GTG

Query Match 0.9%; Score 8; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 679 SSKLVIIIL 686
Db 14 SSKLVIIIL 21

RESULT 11
AD1368
amino acid ABC-transporter (permease) homolog lmo2348 [imported] - *Listeria monocytogenes*
C;Species: *Listeria monocytogenes*
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD1368
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Faihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapk, G.; Madueno, E.; Maitournam, A.; Mak, C.; Schlueter, T.; Smoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AD1368
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-238 <GLA>
A;Cross-references: UNIPROT:Q8YAT4; GB:NC_003210; PIDN:CAD00426.1; PID:gl6411836; GSPDB:AD1368
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2348
C;Superfamily: histidine permease protein M

Query Match 0.9%; Score 8; DB 2; Length 238;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LGVLLSLP 13
Db 38 LGVLLSLP 45

RESULT 12
AF3511
homopermidine synthase [imported] - *Brucella melitensis* (strain 16M)
C;Species: *Brucella melitensis*
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AF3511
R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova, M.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesha, J.; Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*
A;Reference number: AD3252; PMID:11756688
A;Accession: AF3511
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-267 <KUR>
A;Cross-references: UNIPROT:Q8YB02; GB:AE008918; PIDN:AAL53257.1; PID:gl7984137; GSPDB:AF3511
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10016
A;Map position: II

Query Match 0.9%; Score 8; DB 2; Length 267;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 608 VSSAVLAG 615
Db 191 VSSAVLAG 198

RESULT 13
B71686
RNA polymerase sigma-32 factor (rpoH) RP303 - *Rickettsia prowazekii*
C;Species: *Rickettsia prowazekii*
C;Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004
C;Accession: B71686
R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sichert-Ponten, T.; Alsmark, U.; Nature 396, 133-140, 1998
A;Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.
A;Reference number: A71630; MUID:99039499; PMID:9823893
A;Accession: B71686
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-293 <AND>
A;Cross-references: UNIPROT:Q9ZDM4; GB:AJ235271; GB:AJ235269; NID:g3868717; PIDN:CAA1476
A;Experimental source: strain Madrid E
C;Genetics:
A;Gene: rpoH; RP303
C;Superfamily: transcription initiation factor sigma katF; transcription initiation factor sigma katF; transcription initiation factor sigma katF; transcription initiation factor sigma katF
F;55-287/Domain: transcription initiation factor sigma katF homology <KTF>

Query Match 0.9%; Score 8; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 851 SINSDDVE 858
Db 190 SINSDDVE 197

RESULT 14
S61248
hypothetical protein - bovine herpesvirus 1
C;Species: bovine herpesvirus 1
C;Date: 18-Sep-1997 #sequence_revision 18-Sep-1997 #text_change 09-Jul-2004
C;Accession: S61248
R;Vicek, C.; Benes, V.; Lu, Z.; Kutish, G.F.; Paces, V.; Rock, D.; Letchworth, G.J.; Schmitt, J.; Title: Nucleotide sequence analysis of a 30-kb region of the bovine herpesvirus 1 genome.
A;Reference number: S61233
A;Accession: S61248
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <VLC>
A;Cross-references: UNIPROT:Q65574; EMBL:Z48053; NID:g971311; PIDN:CAA88126.1; PID:g97131
C;Superfamily: varicella-zoster virus gene 53 protein

Query Match 0.9%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 870 TFTSSSIT 877
Db 60 TFTSSSIT 67

RESULT 15
A75591
hypothetical protein - *Deinococcus radiodurans* (strain R1)
C;Species: *Deinococcus radiodurans*
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: A75591
R;White, O.; Bisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; S.; Smith, H.O.; Venter, J.C.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.; Shen, M.; Vamathevan, J.J.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266

A;Accession: A75591
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-299 <WHI>
A;Cross-references: UNIPROT:O9RYG0; GB:AE001863; GB:AE001825; NID:G6460670; PIDN:AAF1250
A;Experimental source: strain R1
C;Genetics:
A;Gene: DRA0358
A;Map position: 2

Query Match 0.9%; Score 8; DB 2; Length 299;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 79 SOLLNPL 86
|||
Db 66 SOLLNPL 73

Search completed: July 12, 2005, 07:53:43
Job time : 35 secs

This Page Blank (uspto)

GenCore version 5.1.1.6
 Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:37:57 ; Search time 106 Seconds
 (without alignments)
 4381.659 Million cell updates/sec

Title: US-10-751-736-84

Perfect score: 907

Sequence: 1 MDTSLRLGVLLSLPVLQLAT.....PAYPTVTSCHLSLVAFVPCFL 907

Scoring table: OLIGO
 Gapop 60.0 , Gapext 60.0

Searched: 1612378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : UniProt_03.*
 1: uniprot_sprot.*
 2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	1 LGR5 HUMAN	O75473 homo sapien
2	51	5.6	363	2 Q8C8A7	Q8C8A7 mus musculus
3	51	5.6	907	1 LGR5 MOUSE	Q921P4 mus musculus
4	27	3.0	147	2 Q71M96	Q71M96 sus scrofa
5	18	2.0	828	1 LGR6 HUMAN	Q9HXB8 homo sapien
6	18	2.0	915	2 Q6UY15	Q6UY15 homo sapien
7	18	2.0	923	2 Q86VU0	Q86VU0 homo sapien
8	18	2.0	928	2 Q9BYD7	Q9BYD7 homo sapien
9	15	1.7	878	2 Q8BXS9	Q8BXS9 mus musculus
10	13	1.4	927	2 Q8N537	Q8N537 homo sapien
11	13	1.4	951	1 LGR4 HUMAN	Q9BXB1 homo sapien
12	13	1.4	951	1 LGR4 RAT	Q9Z2H4 rattus norv
13	11	1.2	134	2 Q80T31	Q80T31 mus musculus
14	11	1.2	162	2 Q80UB8	Q80UB8 mus musculus
15	11	1.2	459	2 Q8R301	Q8R301 mus musculus
16	9	1.0	195	2 Q8UH25	Q8UH25 agrobacteri
17	9	1.0	209	2 Q7D1B7	Q7D1B7 agrobacteri
18	9	1.0	369	2 Q93177	Q93177 caenorhabdi
19	9	1.0	379	2 Q2DG05	Q2DG05 gallus gall
20	9	1.0	424	2 Q9DG06	Q9DG06 gallus gall
21	9	1.0	488	2 Q6PHA3	Q6PHA3 mus musculus
22	9	1.0	589	2 Q9FHL8	Q9FHL8 arabidopsis
23	9	1.0	610	2 Q7GEY4	Q7GEY4 podospora a
24	9	1.0	638	2 Q6Z1I6	Q6Z1I6 oryza sativ
25	9	1.0	653	2 Q8LKV9	Q8LKV9 aegilops ta
26	9	1.0	657	2 Q6ZHK7	Q6ZHK7 oryza sativ
27	9	1.0	693	1 F9HR_CHICK	Q79763 gallus gall
28	9	1.0	700	2 Q02714	Q02714 podospora a
29	9	1.0	869	2 Q23253	Q23253 arabidopsis
30	9	1.0	871	2 Q7XNV7	Q7XNV7 oryza sativ
31	9	1.0	879	2 Q68C13	Q68C13 oncorhynch

RESULT 1

LGR5_HUMAN
 ID LGR5_HUMAN STANDARD; PRT; 907 AA.
 AC O75473; OSUP75;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor
 DE (Orphan G protein-coupled receptor HG38) (G protein-coupled receptor 49).
 DE GN Name=GPR49; Synonyms=GPR67, LGR5;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98308104; PubMed=9642114; DOI=10.1006/bbrc.1998.8774;
 RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T., Liu Q.;
 RT "Identification and cloning of an orphan G protein-coupled receptor of the glycoprotein hormone receptor subfamily.";
 RL Biochem. Biophys. Res. Commun. 247:266-270(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;
 RA Hsu S.Y., Liang S.-G., Haueh A.J.W.;
 RT "Characterization of two LGR genes homologous to gonadotropin and thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region.";
 RL Mol. Endocrinol. 12:1830-1845(1998).
 CC -!- FUNCTION: Orphan receptor. It may be an important receptor for embryonic tissues (by similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal cord, and various region of brain.
 CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -!- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.

 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 CC EMBL; AF062006; AAC28019.1; -
 CC EMBL; AF061444; AAC77911.1; -
 CC PIR; JE0176; JE0176.
 CC HSSP; Q9BZR6; IOZN.
 CC Genew; HGNC:4504; GPR49.
 CC MIM; 606667; -.

Q93YT3 arabidopsis
 Q8K7X8 streptococc
 Q8P1F7 streptococc
 Q9A0C0 streptococc
 Q878I7 streptococc
 Q8GUQ5 lycopersico
 Q81899 lycopersico
 Q6L315 solanum dem
 Q6VUQ9 triticum ae
 P73551 synchocyst
 Q91V03 arabidopsis
 Q80A81 ranunculus
 Q863D0 canis famil
 Q863C9 canis famil

DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0004930; P:G-protein coupled receptor activity; TAS.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsin.
 DR InterPro; IPR002131; Gphrmn_receptor.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003372; LRR_Nterm.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00001; 7cm_1; 1.
 DR Pfam; PF00560; LRR; 16.
 DR Pfam; PF01462; LRRNT; 1.
 DR PRINTS; PR00373; GLYCHORMONER.
 DR PRINTS; PR00237; GPCRHHODOPSIN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00369; LRR_TYP; 8.
 DR SMART; SM00013; LRRNT; 1.
 DR PROSITE; PS00237; G-PROTEIN RECP_F1_1; FALSE_NEG.
 DR PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 907 Leucine-rich repeat-containing G protein-coupled receptor 5.
 FT DOMAIN 1 561 Extracellular (Potential).
 FT TRANSMEM 562 582 1 (Potential).
 FT DOMAIN 583 593 Cytoplasmic (Potential).
 FT TRANSMEM 594 614 2 (Potential).
 FT DOMAIN 615 638 Extracellular (Potential).
 FT TRANSMEM 639 659 3 (Potential).
 FT DOMAIN 660 682 Cytoplasmic (Potential).
 FT TRANSMEM 683 703 4 (Potential).
 FT DOMAIN 704 722 Extracellular (Potential).
 FT TRANSMEM 723 743 5 (Potential).
 FT DOMAIN 744 767 Cytoplasmic (Potential).
 FT TRANSMEM 768 788 6 (Potential).
 FT DOMAIN 789 802 Extracellular (Potential).
 FT TRANSMEM 803 823 7 (Potential).
 FT DOMAIN 824 907 Cytoplasmic (Potential).
 FT REPEAT 64 88 LRR 1.
 FT REPEAT 89 112 LRR 2.
 FT REPEAT 114 136 LRR 3.
 FT REPEAT 137 160 LRR 4.
 FT REPEAT 162 184 LRR 5.
 FT REPEAT 185 208 LRR 6.
 FT REPEAT 209 232 LRR 7.
 FT REPEAT 233 256 LRR 8.
 FT REPEAT 257 279 LRR 9.
 FT REPEAT 281 303 LRR 10.
 FT REPEAT 304 327 LRR 11.
 FT REPEAT 328 350 LRR 12.
 FT REPEAT 351 375 LRR 13.
 FT REPEAT 377 396 LRR 14.
 FT REPEAT 397 420 LRR 15.
 FT REPEAT 422 444 LRR 16.
 FT REPEAT 445 468 LRR 17.
 FT REPEAT 469 492 LRR 18.
 FT REPEAT 493 516 LRR 19.
 FT REPEAT 517 540 LRR 20.
 FT REPEAT 541 564 LRR 21.
 FT REPEAT 565 588 LRR 22.
 FT REPEAT 589 612 LRR 23.
 FT REPEAT 613 636 LRR 24.
 FT REPEAT 637 660 LRR 25.
 FT REPEAT 661 684 LRR 26.
 FT REPEAT 685 708 LRR 27.
 FT REPEAT 709 732 LRR 28.
 FT REPEAT 733 756 LRR 29.
 FT REPEAT 757 780 LRR 30.
 FT REPEAT 781 804 LRR 31.
 FT REPEAT 805 828 LRR 32.
 FT REPEAT 829 852 LRR 33.
 FT REPEAT 853 876 LRR 34.
 FT REPEAT 877 900 LRR 35.
 FT REPEAT 901 924 LRR 36.
 FT REPEAT 925 948 LRR 37.
 FT REPEAT 949 972 LRR 38.
 FT REPEAT 973 996 LRR 39.
 FT REPEAT 997 1020 LRR 40.
 FT REPEAT 1021 1044 LRR 41.
 FT REPEAT 1045 1068 LRR 42.
 FT REPEAT 1069 1092 LRR 43.
 FT REPEAT 1093 1116 LRR 44.
 FT REPEAT 1117 1140 LRR 45.
 FT REPEAT 1141 1164 LRR 46.
 FT REPEAT 1165 1188 LRR 47.
 FT REPEAT 1189 1212 LRR 48.
 FT REPEAT 1213 1236 LRR 49.
 FT REPEAT 1237 1260 LRR 50.
 FT REPEAT 1261 1284 LRR 51.
 FT REPEAT 1285 1308 LRR 52.
 FT REPEAT 1309 1332 LRR 53.
 FT REPEAT 1333 1356 LRR 54.
 FT REPEAT 1357 1380 LRR 55.
 FT REPEAT 1381 1404 LRR 56.
 FT REPEAT 1405 1428 LRR 57.
 FT REPEAT 1429 1452 LRR 58.
 FT REPEAT 1453 1476 LRR 59.
 FT REPEAT 1477 1500 LRR 60.
 FT REPEAT 1501 1524 LRR 61.
 FT REPEAT 1525 1548 LRR 62.
 FT REPEAT 1549 1572 LRR 63.
 FT REPEAT 1573 1596 LRR 64.
 FT REPEAT 1597 1620 LRR 65.
 FT REPEAT 1621 1644 LRR 66.
 FT REPEAT 1645 1668 LRR 67.
 FT REPEAT 1669 1692 LRR 68.
 FT REPEAT 1693 1716 LRR 69.
 FT REPEAT 1717 1740 LRR 70.
 FT REPEAT 1741 1764 LRR 71.
 FT REPEAT 1765 1788 LRR 72.
 FT REPEAT 1789 1812 LRR 73.
 FT REPEAT 1813 1836 LRR 74.
 FT REPEAT 1837 1860 LRR 75.
 FT REPEAT 1861 1884 LRR 76.
 FT REPEAT 1885 1908 LRR 77.
 FT REPEAT 1909 1932 LRR 78.
 FT REPEAT 1933 1956 LRR 79.
 FT REPEAT 1957 1980 LRR 80.
 FT REPEAT 1981 2004 LRR 81.
 FT REPEAT 2005 2028 LRR 82.
 FT REPEAT 2029 2052 LRR 83.
 FT REPEAT 2053 2076 LRR 84.
 FT REPEAT 2077 2100 LRR 85.
 FT REPEAT 2101 2124 LRR 86.
 FT REPEAT 2125 2148 LRR 87.
 FT REPEAT 2149 2172 LRR 88.
 FT REPEAT 2173 2196 LRR 89.
 FT REPEAT 2197 2220 LRR 90.
 FT REPEAT 2221 2244 LRR 91.
 FT REPEAT 2245 2268 LRR 92.
 FT REPEAT 2269 2292 LRR 93.
 FT REPEAT 2293 2316 LRR 94.
 FT REPEAT 2317 2340 LRR 95.
 FT REPEAT 2341 2364 LRR 96.
 FT REPEAT 2365 2388 LRR 97.
 FT REPEAT 2389 2412 LRR 98.
 FT REPEAT 2413 2436 LRR 99.
 FT REPEAT 2437 2460 LRR 100.
 FT REPEAT 2461 2484 LRR 101.
 FT REPEAT 2485 2508 LRR 102.
 FT REPEAT 2509 2532 LRR 103.
 FT REPEAT 2533 2556 LRR 104.
 FT REPEAT 2557 2580 LRR 105.
 FT REPEAT 2581 2604 LRR 106.
 FT REPEAT 2605 2628 LRR 107.
 FT REPEAT 2629 2652 LRR 108.
 FT REPEAT 2653 2676 LRR 109.
 FT REPEAT 2677 2700 LRR 110.
 FT REPEAT 2701 2724 LRR 111.
 FT REPEAT 2725 2748 LRR 112.
 FT REPEAT 2749 2772 LRR 113.
 FT REPEAT 2773 2796 LRR 114.
 FT REPEAT 2797 2820 LRR 115.
 FT REPEAT 2821 2844 LRR 116.
 FT REPEAT 2845 2868 LRR 117.
 FT REPEAT 2869 2892 LRR 118.
 FT REPEAT 2893 2916 LRR 119.
 FT REPEAT 2917 2940 LRR 120.
 FT REPEAT 2941 2964 LRR 121.
 FT REPEAT 2965 2988 LRR 122.
 FT REPEAT 2989 3012 LRR 123.
 FT REPEAT 3013 3036 LRR 124.
 FT REPEAT 3037 3060 LRR 125.
 FT REPEAT 3061 3084 LRR 126.
 FT REPEAT 3085 3108 LRR 127.
 FT REPEAT 3109 3132 LRR 128.
 FT REPEAT 3133 3156 LRR 129.
 FT REPEAT 3157 3180 LRR 130.
 FT REPEAT 3181 3204 LRR 131.
 FT REPEAT 3205 3228 LRR 132.
 FT REPEAT 3229 3252 LRR 133.
 FT REPEAT 3253 3276 LRR 134.
 FT REPEAT 3277 3300 LRR 135.
 FT REPEAT 3301 3324 LRR 136.
 FT REPEAT 3325 3348 LRR 137.
 FT REPEAT 3349 3372 LRR 138.
 FT REPEAT 3373 3396 LRR 139.
 FT REPEAT 3397 3420 LRR 140.
 FT REPEAT 3421 3444 LRR 141.
 FT REPEAT 3445 3468 LRR 142.
 FT REPEAT 3469 3492 LRR 143.
 FT REPEAT 3493 3516 LRR 144.
 FT REPEAT 3517 3540 LRR 145.
 FT REPEAT 3541 3564 LRR 146.
 FT REPEAT 3565 3588 LRR 147.
 FT REPEAT 3589 3612 LRR 148.
 FT REPEAT 3613 3636 LRR 149.
 FT REPEAT 3637 3660 LRR 150.
 FT REPEAT 3661 3684 LRR 151.
 FT REPEAT 3685 3708 LRR 152.
 FT REPEAT 3709 3732 LRR 153.
 FT REPEAT 3733 3756 LRR 154.
 FT REPEAT 3757 3780 LRR 155.
 FT REPEAT 3781 3804 LRR 156.
 FT REPEAT 3805 3828 LRR 157.
 FT REPEAT 3829 3852 LRR 158.
 FT REPEAT 3853 3876 LRR 159.
 FT REPEAT 3877 3900 LRR 160.
 FT REPEAT 3901 3924 LRR 161.
 FT REPEAT 3925 3948 LRR 162.
 FT REPEAT 3949 3972 LRR 163.
 FT REPEAT 3973 3996 LRR 164.
 FT REPEAT 3997 4020 LRR 165.
 FT REPEAT 4021 4044 LRR 166.
 FT REPEAT 4045 4068 LRR 167.
 FT REPEAT 4069 4092 LRR 168.
 FT REPEAT 4093 4116 LRR 169.
 FT REPEAT 4117 4140 LRR 170.
 FT REPEAT 4141 4164 LRR 171.
 FT REPEAT 4165 4188 LRR 172.
 FT REPEAT 4189 4212 LRR 173.
 FT REPEAT 4213 4236 LRR 174.
 FT REPEAT 4237 4260 LRR 175.
 FT REPEAT 4261 4284 LRR 176.
 FT REPEAT 4285 4308 LRR 177.
 FT REPEAT 4309 4332 LRR 178.
 FT REPEAT 4333 4356 LRR 179.
 FT REPEAT 4357 4380 LRR 180.
 FT REPEAT 4381 4404 LRR 181.
 FT REPEAT 4405 4428 LRR 182.
 FT REPEAT 4429 4452 LRR 183.
 FT REPEAT 4453 4476 LRR 184.
 FT REPEAT 4477 4500 LRR 185.
 FT REPEAT 4501 4524 LRR 186.
 FT REPEAT 4525 4548 LRR 187.
 FT REPEAT 4549 4572 LRR 188.
 FT REPEAT 4573 4596 LRR 189.
 FT REPEAT 4597 4620 LRR 190.
 FT REPEAT 4621 4644 LRR 191.
 FT REPEAT 4645 4668 LRR 192.
 FT REPEAT 4669 4692 LRR 193.
 FT REPEAT 4693 4716 LRR 194.
 FT REPEAT 4717 4740 LRR 195.
 FT REPEAT 4741 4764 LRR 196.
 FT REPEAT 4765 4788 LRR 197.
 FT REPEAT 4789 4812 LRR 198.
 FT REPEAT 4813 4836 LRR 199.
 FT REPEAT 4837 4860 LRR 200.
 FT REPEAT 4861 4884 LRR 201.
 FT REPEAT 4885 4908 LRR 202.
 FT REPEAT 4909 4932 LRR 203.
 FT REPEAT 4933 4956 LRR 204.
 FT REPEAT 4957 4980 LRR 205.
 FT REPEAT 4981 5004 LRR 206.
 FT REPEAT 5005 5028 LRR 207.
 FT REPEAT 5029 5052 LRR 208.
 FT REPEAT 5053 5076 LRR 209.
 FT REPEAT 5077 5100 LRR 210.
 FT REPEAT 5101 5124 LRR 211.
 FT REPEAT 5125 5148 LRR 212.
 FT REPEAT 5149 5172 LRR 213.
 FT REPEAT 5173 5196 LRR 214.
 FT REPEAT 5197 5220 LRR 215.
 FT REPEAT 5221 5244 LRR 216.
 FT REPEAT 5245 5268 LRR 217.
 FT REPEAT 5269 5292 LRR 218.
 FT REPEAT 5293 5316 LRR 219.
 FT REPEAT 5317 5340 LRR 220.
 FT REPEAT 5341 5364 LRR 221.
 FT REPEAT 5365 5388 LRR 222.
 FT REPEAT 5389 5412 LRR 223.
 FT REPEAT 5413 5436 LRR 224.
 FT REPEAT 5437 5460 LRR 225.
 FT REPEAT 5461 5484 LRR 226.
 FT REPEAT 5485 5508 LRR 227.
 FT REPEAT 5509 5532 LRR 228.
 FT REPEAT 5533 5556 LRR 229.
 FT REPEAT 5557 5580 LRR 230.
 FT REPEAT 5581 5604 LRR 231.
 FT REPEAT 5605 5628 LRR 232.
 FT REPEAT 5629 5652 LRR 233.
 FT REPEAT 5653 5676 LRR 234.
 FT REPEAT 5677 5700 LRR 235.
 FT REPEAT 5701 5724 LRR 236.
 FT REPEAT 5725 5748 LRR 237.
 FT REPEAT 5749 5772 LRR 238.
 FT REPEAT 5773 5796 LRR 239.
 FT REPEAT 5797 5820 LRR 240.
 FT REPEAT 5821 5844 LRR 241.
 FT REPEAT 5845 5868 LRR 242.
 FT REPEAT 5869 5892 LRR 243.
 FT REPEAT 5893 5916 LRR 244.
 FT REPEAT 5917 5940 LRR 245.
 FT REPEAT 5941 5964 LRR 246.
 FT REPEAT 5965 5988 LRR 247.
 FT REPEAT 5989 6012 LRR 248.
 FT REPEAT 6013 6036 LRR 249.
 FT REPEAT 6037 6060 LRR 250.
 FT REPEAT 6061 6084 LRR 251.
 FT REPEAT 6085 6108 LRR 252.
 FT REPEAT 6109 6132 LRR 253.
 FT REPEAT 6133 6156 LRR 254.
 FT REPEAT 6157 6180 LRR 255.
 FT REPEAT 6181 6204 LRR 256.
 FT REPEAT 6205 6228 LRR 257.
 FT REPEAT 6229 6252 LRR 258.
 FT REPEAT 6253 6276 LRR 259.
 FT REPEAT 6277 6300 LRR 260.
 FT REPEAT 6301 6324 LRR 261.
 FT REPEAT 6325 6348 LRR 262.
 FT REPEAT 6349 6372 LRR 263.
 FT REPEAT 6373 6396 LRR 264.
 FT REPEAT 6397 6420 LRR 265.
 FT REPEAT 6421 6444 LRR 266.
 FT REPEAT 6445 6468 LRR 267.
 FT REPEAT 6469 6492 LRR 268.
 FT REPEAT 6493 6516 LRR 269.
 FT REPEAT 6517 6540 LRR 270.
 FT REPEAT 6541 6564 LRR 271.
 FT REPEAT 6565 6588 LRR 272.
 FT REPEAT 6589 6612 LRR 273.
 FT REPEAT 6613 6636 LRR 274.
 FT REPEAT 6637 6660 LRR 275.
 FT REPEAT 6661 6684 LRR 276.
 FT REPEAT 6685 6708 LRR 277.
 FT REPEAT 6709 6732 LRR 278.
 FT REPEAT 6733 6756 LRR 279.
 FT REPEAT 6757 6780 LRR 280.
 FT REPEAT 6781 6804 LRR 281.
 FT REPEAT 6805 6828 LRR 282.
 FT REPEAT 6829 6852 LRR 283.
 FT REPEAT 6853 6876 LRR 284.
 FT REPEAT 6877 6900 LRR 285.
 FT REPEAT 6901 6924 LRR 286.
 FT REPEAT 6925 6948 LRR 287.
 FT REPEAT 6949 6972 LRR 288.
 FT REPEAT 6973 6996 LRR 289.
 FT REPEAT 6997 7020 LRR 290.
 FT REPEAT 7021 7044 LRR 291.
 FT REPEAT 7045 7068 LRR 292.
 FT REPEAT 7069 7092 LRR 293.
 FT REPEAT 7093 7116 LRR 294.
 FT REPEAT 7117 7140 LRR 295.
 FT REPEAT 7141 7164 LRR 296.
 FT REPEAT 7165 7188 LRR 297.
 FT REPEAT 7189 7212 LRR 298.
 FT REPEAT 7213 7236 LRR 299.
 FT REPEAT 7237 7260 LRR 300.
 FT REPEAT 7261 7284 LRR 301.
 FT REPEAT 7285 7308 LRR 302.
 FT REPEAT 7309 7332 LRR 303.
 FT REPEAT 7333 7356 LRR 304.
 FT REPEAT 7357 7380 LRR 305.
 FT REPEAT 7381 7404 LRR 306.
 FT REPEAT 7405 7428 LRR 307.
 FT REPEAT 7429 7452 LRR 308.
 FT REPEAT 7453 7476 LRR 309.
 FT REPEAT 7477 7500 LRR 310.
 FT REPEAT 7501 7524 LRR 311.
 FT REPEAT 7525 7548 LRR 312.
 FT REPEAT 7549 7572 LRR 313.
 FT REPEAT 7573 7596 LRR 314.
 FT REPEAT 7597 7620 LRR 315.
 FT REPEAT 7621 7644 LRR 316.
 FT REPEAT 7645 7668 LRR 317.
 FT REPEAT 7669 7692 LRR 318.
 FT REPEAT 7693 7716 LRR 319.
 FT REPEAT 7717 7740 LRR 320.
 FT REPEAT 7741 7764 LRR 321.
 FT REPEAT 7765 7788 LRR 322.
 FT REPEAT 7789 7812 LRR 323.
 FT REPEAT 7813 7836 LRR 324.
 FT REPEAT 7837 7860 LRR 325.
 FT REPEAT 7861 7884 LRR 326.
 FT REPEAT 7885 7908 LRR 327.
 FT REPEAT 7909 7932 LRR 328.
 FT REPEAT 7933 7956 LRR 329.
 FT REPEAT 7957 7980 LRR 330.
 FT REPEAT 7981 8004 LRR 331.
 FT REPEAT 8005 8028 LRR 332.
 FT REPEAT 8029 8052 LRR 333.
 FT REPEAT 8053 8076 LRR 334.
 FT REPEAT 8077 8100 LRR 335.
 FT REPEAT 8101 8124 LRR 336.
 FT REPEAT 8125 8148 LRR 337.
 FT REPEAT 8149 8172 LRR 338.
 FT REPEAT 8173 8196 LRR 339.
 FT REPEAT 8197 8220 LRR 340.
 FT REPEAT 8221 8244 LRR 341.
 FT REPEAT 8245 8268 LRR 342.
 FT REPEAT 8269 8292 LRR 343.
 FT REPEAT 8293 8316 LRR 344.
 FT REPEAT 8317 8340 LRR 345.
 FT REPEAT 8341 8364 LRR 346.
 FT REPEAT 8365 8388 LRR 347.
 FT REPEAT 8389 8412 LRR 348.
 FT REPEAT 8413 8436 LRR 349.
 FT REPEAT 8437 8460 LRR 350.
 FT REPEAT 8461 8484 LRR 351.
 FT REPEAT 8485 8508 LRR 352.
 FT REPEAT 8509 8532 LRR 353.
 FT REPEAT 8533 8556 LRR 354.
 FT REPEAT 8557 8580 LRR 355.
 FT REPEAT 8581 8604 LRR 356.
 FT REPEAT 8605 8628 LRR 357.
 FT REPEAT 8629 8652 LRR 358.
 FT REPEAT 8653 8676 LRR 359.
 FT REPEAT 8677 8700 LRR 360.
 FT REPEAT 8701 8724 LRR 361.
 FT REPEAT 8725 8748 LRR 362.
 FT REPEAT 8749 8772 LRR 363.
 FT REPEAT 8773 8796 LRR 364.
 FT REPEAT 8797 8820 LRR 365.
 FT REPEAT 8821 8844 LRR 366.
 FT REPEAT 8845 8868 LRR 367.
 FT REPEAT 8869 8892 LRR 368.
 FT REPEAT 8893 8916 LRR 369.
 FT REPEAT 8917 8940 LRR 370.
 FT REPEAT 8941 8964 LRR 371.
 FT REPEAT 8965 8988 LRR 372.
 FT REPEAT 8989 9012 LRR 373.
 FT REPEAT 9013 9036 LRR 374.
 FT REPEAT 9037 9060 LRR 375.
 FT REPEAT 9061 9084 LRR 376.
 FT REPEAT 9085 9108 LRR 377.
 FT REPEAT 9109 9132 LRR 378.
 FT REPEAT 9133 9156 LRR 379.
 FT REPEAT 9157 9180 LRR 380.
 FT REPEAT 9181 9204 LRR 381.
 FT REPEAT 9205 9228 LRR 382.
 FT REPEAT 9229 9252 LRR 383.
 FT REPEAT 9253 9276 LRR 384.
 FT REPEAT 9277 9300 LRR 385.
 FT REPEAT 9301 9324 LRR 386.
 FT REPEAT 9325 9348 LRR 387.
 FT REPEAT 9349 9372 LRR 388.
 FT REPEAT 9373 9396 LRR 389.
 FT REPEAT 9397 9420 LRR 390.
 FT REPEAT 9421 9444 LRR 391.
 FT REPEAT 9445 9468 LRR 392.
 FT REPEAT 9469 9492 LRR 393.
 FT REPEAT 9493 9516 LRR 394.
 FT REPEAT 9517 9540 LRR 395.
 FT REPEAT 9541 9564 LRR 396.
 FT REPEAT 9565 9588 LRR 397.
 FT REPEAT 9589 9612 LRR 398.
 FT REPEAT 9613 9636 LRR 399.
 FT REPEAT 9637 9660 LRR 400.
 FT REPEAT 9661 9684 LRR 401.
 FT REPEAT 9685 9708 LRR 402.
 FT REPEAT 9709 9732 LRR 403.
 FT REPEAT 9733 9756 LRR 404.
 FT REPEAT 9757 9780 LRR 405.
 FT REPEAT 9781 9804 LRR 406.
 FT REPEAT 9805 9828 LRR 407.
 FT REPEAT 9829 9852 LRR 408.
 FT REPEAT 9853 9876 LRR 409.
 FT REPEAT 9877 9900 LRR 410.
 FT REPEAT 9901 9924 LRR 411.
 FT REPEAT 9925 9948 LRR 412.
 FT REPEAT 9949 9972 LRR 413.
 FT REPEAT 9973 10000 LRR 414.
 FT REPEAT 10001 10024 LRR 415.
 FT REPEAT 10025 10048 LRR 416.
 FT REPEAT 10049 10072 LRR 417.
 FT REPEAT 10073 10096 LRR 418.
 FT REPEAT 10097 10120 LRR 419.
 FT REPEAT 10121 10144 LRR 420.
 FT REPEAT 10145 10168 LRR 421.
 FT REPEAT 10169 10192 LRR 422.
 FT REPEAT 10193 10216 LRR 423.
 FT REPEAT 10217 10240 LRR 424.
 FT REPEAT 10241 10264 LRR 425.
 FT REPEAT 10265 10288 LRR 426.
 FT REPEAT 10289 10312 LRR 427.
 FT REPEAT 10313 10336 LRR 428.
 FT REPEAT 10337 10360 LRR 429.
 FT REPEAT 10361 10384 LRR 430.
 FT REPEAT 10385 10408 LRR 431.
 FT REPEAT 10409 10432 LRR 432.
 FT REPEAT 10433 10456 LRR 433.
 FT REPEAT 10457 10480 LRR 434.
 FT REPEAT 10481 10504 LRR 435.
 FT REPEAT 10505 10528 LRR 436.
 FT REPEAT 10529 10552 LRR 437.
 FT REPEAT 10553 10576 LRR 438.
 FT REPEAT 10577 10600 LRR 439.
 FT REPEAT 10601 10624 LRR 440.
 FT REPEAT 10625 10648 LRR 441.
 FT REPEAT 10649 10672 LRR 442.
 FT REPEAT 10673 10696 LRR 443.
 FT REPEAT 10697 10720 LRR 444.
 FT REPEAT 10721 10744 LRR 445.
 FT REPEAT 10745 10768 LRR 446.
 FT REPEAT 10769 10792 LRR 447.
 FT REPEAT 10793 10816 LRR 448.
 FT REPEAT 10817 10840 LRR 449.
 FT REPEAT 10841 10864 LRR 450.
 FT REPEAT 10865 10888 LRR 451.
 FT REPEAT 10889 10912 LRR 452.
 FT REPEAT 10913 10936 LRR 453.
 FT REPEAT 10937 10960 LRR 454.
 FT REPEAT 10961 10984 LRR 455.
 FT REPEAT 10985 11008 LRR 456.
 FT REPEAT 11009 11032 LRR 457.
 FT REPEAT 11033 11056 LRR 458.
 FT REPEAT 11057 11080 LRR 459.
 FT REPEAT 11081 11104 LRR 460.
 FT REPEAT 11105 11128 LRR 461.
 FT REPEAT 11129 11152 LRR 462.
 FT REPEAT 11153 11176 LRR 463.
 FT REPEAT 11177 11200 LRR 464.
 FT REPEAT 11201 11224 LRR 465.
 FT REPEAT 11225 11248 LRR 466.
 FT REPEAT 11249 11272 LRR 467.
 FT REPEAT 11273 11296 LRR 468.
 FT REPEAT 11297 11320 LRR 469.
 FT REPEAT 11321 11344 LRR 470.
 FT REPEAT 11345 11368 LRR 471.
 FT REPEAT 11369 11392 LRR 472.
 FT REPEAT 11393 11416 LRR 473.
 FT REPEAT 11417 11440 LRR 474.
 FT REPEAT 11441 11464 LRR 475.
 FT REPEAT 11465 11488 LRR 476.
 FT REPEAT 11489 11512 LRR 477.
 FT REPEAT 11513


```

DR Pfam; PF00560; LRR; 16.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHOPOFSN.
DR PRINTS; SM00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYF; 8.
DR SMART; SM00013; LRRNT; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 907
FT DOMAIN 22 561
FT TRANSMEM 562 582
FT DOMAIN 583 593
FT TRANSMEM 594 614
FT DOMAIN 615 638
FT TRANSMEM 639 659
FT DOMAIN 660 682
FT TRANSMEM 683 703
FT DOMAIN 704 723
FT TRANSMEM 724 744
FT DOMAIN 745 767
FT TRANSMEM 768 788
FT DOMAIN 789 802
FT TRANSMEM 803 823
FT DOMAIN 824 907
FT REPEAT 64 88
FT REPEAT 89 112
FT REPEAT 113 136
FT REPEAT 137 160
FT REPEAT 162 184
FT REPEAT 186 208
FT REPEAT 209 232
FT REPEAT 233 256
FT REPEAT 257 279
FT REPEAT 281 303
FT REPEAT 304 327
FT REPEAT 328 350
FT REPEAT 351 375
FT REPEAT 377 396
FT REPEAT 397 420
FT REPEAT 422 444
FT REPEAT 447 470
FT CARBOHYD 63 63 N-linked (GlcNAc...) (potential).
FT CARBOHYD 77 77 N-linked (GlcNAc...) (potential).
FT CARBOHYD 208 208 N-linked (GlcNAc...) (potential).
FT CARBOHYD 792 792 N-linked (GlcNAc...) (potential).
SQ SEQUENCE 907 AA; 99681 MW; 553167C6C0AAE253 CRC64;

Query Match 5.6%; Score 51; DB 1; Length 907;
Best Local Similarity 100.0%; Pred. No. 2.1e-39;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 203 DYAFGNLSLVVLHNNRIHSLGKCPDGLHSLTLDLNNLDEPPTAI 253
Db 203 DYAFGNLSLVVLHNNRIHSLGKCPDGLHSLTLDLNNLDEPPTAI 253

RESULT 4
ID Q71M96 PRELIMINARY; PRT; 147 AA.
AC Q71M96;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE G-protein-coupled receptor 49 (fragment).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
NCBI_TaxID=9823;

```

```

RN RP SEQUENCE FROM N.A.
RA Malek M., Ciobanu D.C., Rothschild M.F.; DDBJ databases.
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF455789; AAQ04721.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0004872; P:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin.; IEA.
DR InterPro; IPR000276; GPCR Rhodopsin.
DR InterPro; IPR002131; Gphrmn receptor.
DR PRINTS; PR00373; GLYCHORMONER.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 147 147
SQ SEQUENCE 147 AA; 15250 MW; 01AE345327EB3586 CRC64;

Query Match 3.0%; Score 27; DB 2; Length 147;
Best Local Similarity 100.0%; Pred. No. 5.6e-17;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 641 GFLSIFASESSVFLTLAALERGFVSX 667
Db 78 GFLSIFASESSVFLTLAALERGFVSX 104

RESULT 5
LGR6 HUMAN
ID LGR6 HUMAN STANDARD; PRT; 828 AA.
AC Q9HBX8; O96K69;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 6.
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20388592; PubMed=10935549; DOI=10.1210/me.14.8.1257;
RA Hsu S.Y., Kudo M., Chen T., Nakabayashi K., Bhalla A.,
RA van der Spek P.J., van Duin M., Hsueh A.J.W.;
RT "The three subfamilies of leucine-rich repeat-containing G protein-
RT coupled receptors (LGR): identification of LGR6 and LGR7 and the
RT signaling mechanism for LGR7."
RL Mol. Endocrinol. 14:1257-1271(2000).
RN [2]
RP SEQUENCE OF 406-828 FROM N.A.
RX MEDLINE=22040266; PubMed=12044878; DOI=10.1016/S0014-5793(02)02775-8;
RA Takeda S., Kadowaki S., Haga T., Takaue H., Mitaku S.;
RT "Identification of G protein-coupled receptor genes from the human
RT genome sequence."
RL FEBS Lett. 520:97-101(2002).
RN [3]
RP SEQUENCE OF 431-828 FROM N.A.
RC TISSUE=Mammary gland;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahari K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa K.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoaka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotoa T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K., Arita M.,

```


RA Imose N., Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
 RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
 RA Moriya S., Momiyama H., Satcho N., Takami N., Terashima Y., Suzuki O.,
 RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
 RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamagaki M., Watanabe K., Kamagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiya M., Tashiro H., Tanigami A., Fujiwara T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Omori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
 RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satcho T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Negase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara G., Isogai T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 CC -!- FUNCTION: Orphan receptor.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -!- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF190501; AAG17168.1; -.
 DR EMBL; AB083616; BAB89329.1; -.
 DR EMBL; AK027377; BAB55071.1; ALT_INIT.
 DR HSSP; P25147; 1DOB.
 DR Genew; HGNC:19719; LGR6.
 DR MIM; 606653; -.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR002131; Gphrmn_receptor.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00560; LRR; 12.
 DR PRINTS; PR00237; GLYCHORMONER.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PRINTS; PR00119; LEURICHRPT.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; FALSE_NEG.
 DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; FALSE_NEG.
 KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
 Transmembrane.
 FT DOMAIN 1 428 Extracellular (Potential).
 FT TRANSMEM 429 449 1 (Potential).
 FT DOMAIN 450 459 Cytoplasmic (Potential).
 FT TRANSMEM 460 480 2 (Potential).
 FT DOMAIN 481 505 Extracellular (Potential).
 FT TRANSMEM 506 526 3 (Potential).
 FT DOMAIN 527 548 Cytoplasmic (Potential).
 FT TRANSMEM 549 569 4 (Potential).
 FT DOMAIN 570 588 Extracellular (Potential).
 FT TRANSMEM 589 609 5 (Potential).
 FT DOMAIN 610 635 Cytoplasmic (Potential).
 FT TRANSMEM 636 656 6 (Potential).
 FT DOMAIN 657 670 Extracellular (Potential).
 FT TRANSMEM 671 691 7 (Potential).
 FT DOMAIN 692 804 Cytoplasmic (Potential).
 FT TRANSMEM 805 816 Poly-Gly.
 FT REPEAT 22 45 LRR 1.
 FT REPEAT 46 69 LRR 2.
 FT REPEAT 71 93 LRR 3.
 FT REPEAT 94 117 LRR 4.
 FT REPEAT 118 140 LRR 5.
 FT REPEAT 142 164 LRR 6.
 FT REPEAT 188 211 LRR 7.
 FT REPEAT 212 236 LRR 8.

FT REPEAT 238 257 LRR 9.
 FT REPEAT 258 281 LRR 10.
 FT REPEAT 283 305 LRR 11.
 FT DISULFID 503 578 By similarity.
 FT CARBOHYD 15 15 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 34 34 N-linked (GlcNAc...) (Potential).
 FT CONFLICT 406 410 CSPTP -> MISPT (in Ref. 2).
 FT CONFLICT 628 628 W -> R (in Ref. 3).
 FT CONFLICT 824 828 FASHV -> LLHTY (in Ref. 1).
 SQ SEQUENCE 828 AA; 89301 MW; 1B5971445AA2D8B4 CRC64;
 Query Match 2.0%; Score 18; DB 1; Length 828;
 Best Local Similarity 100.0%; Pred. No. 1.1e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 285 TIHFYDNPQFVGRSAFQ 302
 DB 146 TIHFYDNPQFVGRSAFQ 163
 RESULT 6
 Q6UY15 PRELIMINARY; PRT; 915 AA.
 AC Q6UY15
 DT 05-JUL-2004 (TRENBLrel. 27, Created)
 DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
 GE Gonadotropin receptor.
 GN ORFNames=UNQ6427.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22897296; PubMed=12975309; DOI=10.1101/gr.1293003;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heldens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R., Watanabe C., Wiedand D., Woods K., Xie M.H., Yansura D.,
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
 RA Godowski P.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270(2003).
 DR EMBL; AY358119; AAQ88486.1; -.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR002131; Gphrmn_receptor.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR003885; LRR_cyst.
 DR InterPro; IPR003591; LRR_typ.
 DR Pfam; PF00560; LRR_1; 15.
 DR PRINTS; PR00373; GLYCHORMONER.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PRINTS; PR00119; LEURICHRPT.
 DR SMART; SM00365; LRR_SD22; 5.
 DR SMART; SM00369; LRR_TYP; 14.
 KW Receptor.
 SQ SEQUENCE 915 AA; 99265 MW; D57DD7A9DBB555F4 CRC64;
 Query Match 2.0%; Score 18; DB 2; Length 915;
 Best Local Similarity 100.0%; Pred. No. 1.2e-07;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 285 TIHFYDNPQFVGRSAFQ 302


```
Db 233 TIHFYDNPQFVGRSAFQ 250
|||||
RESULT 7
ID Q86VU0 PRELIMINARY; PRT; 923 AA.
AC Q86VU0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LGR6 protein (Fragment).
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=2328257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC047905; AAH47905.1; -.
DR HSSP; Q9BZR6; 1P8T.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gphrmn_receptor.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.
FT NON_TER 1
SQ SEQUENCE 923 AA; 10047 MW; 541D6746DAB06813 CRC64;
Query Match 2.0%; Score 18; DB 2; Length 923;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 285 TIHFYDNPQFVGRSAFQ 302
|||||
Db 241 TIHFYDNPQFVGRSAFQ 258
|||||
RESULT 8
QY 285 TIHFYDNPQFVGRSAFQ 302
|||||
Db 241 TIHFYDNPQFVGRSAFQ 258
|||||
Q9BYD7 PRELIMINARY; PRT; 928 AA.
AC Q9BYD7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE VTS20631 (Fragment).
GN Name=VTS20631;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Okaze H., Hayashi A., Kozuma S., Saito T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AB049405; BAB39854.1; -.
DR HSSP; P25147; 1D0B.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gphrmn_receptor.
DR InterPro; IPR003591; LRR_Typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.
FT NON_TER 1
SQ SEQUENCE 928 AA; 100487 MW; 4C3364ADEA89C463 CRC64;
Query Match 2.0%; Score 18; DB 2; Length 928;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 285 TIHFYDNPQFVGRSAFQ 302
|||||
Db 246 TIHFYDNPQFVGRSAFQ 263
|||||
RESULT 9
Q8BXS9 PRELIMINARY; PRT; 878 AA.
AC Q8BXS9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus adult retina cDNA, RIKEN full-length enriched library,
clone:A930003A08 product:G PROTEIN-COUPLED RECEPTOR LGRA homolog
(DE Fragment).
DE (Fragment).
GN Name=Gpr48;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
```


RC STRAIN=C57BL/6J; TISSUE=Retina;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RN Nature 420:563-573 (2002).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata K., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630 (2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakauchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohata E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
[6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi P., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK044357; BAC31882.1; -
DR HSSP; P25147; 1D08.
DR MGD; MGI:1891468; Gpr48.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002131; Gprmrn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_Nterm.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00001; 7tm 1; 1.
DR Pfam; PF00560; LRR 1; 14.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Receptor.
FT NON_TER.
FT SEQUENCE 878 AA; 96877 MW; 68E64B5E5DEA11B37 CRC64;
Query Match 1.7%; Score 15; DB 2; Length 878;
Best Local Similarity 100.0%; Pred. No. 8.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 236 LETLDLNNYNNLDEFP 250

Db 154 LETLDLNNYNNLDEFP 168
RESULT 10
Q8N537 PRELIMINARY; PRT; 927 AA.
ID Q8N537
AC Q8N537
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE GPR48 protein.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.B., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC033039; AAH33039.1; -
DR HSSP; P25147; 1D08.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002131; Gprmrn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_Typ.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00001; 7tm 1; 1.
DR Pfam; PF01462; LRRNT; 1.
DR Pfam; PF00560; LRR 1; 14.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHOPOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SQ SEQUENCE 927 AA; 101676 MW; C7E2F0C40E584C8 CRC64;
Query Match 1.4%; Score 13; DB 2; Length 927;
Best Local Similarity 100.0%; Pred. No. 0.0076;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 208 NLSSLVVLHNN 220


```

Db      175 NLSSLVVLHNN 187
|||||
ID_LGR4_HUMAN STANDARD; PRT; 951 AA.
AC Q9XB1; Q9YD1;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
DE (G protein-coupled receptor 48).
GN Name=GPR48; Synonyms=LGR4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreas;
RX MEDLINE=21294803; PubMed=11401528; DOI=10.1006/bbrc.2001.4625;
RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
RT "Molecular characterization of a novel glycoprotein hormone G-protein-
coupled receptor.";
RL Biochem. Biophys. Res. Commun. 282:757-764(2001).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues:
placenta, ovary, testis and adrenal. Expressed also in spinal
cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate
and spleen.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AF346711; AAK31153.1; -
EMBL; AF346709; AAK31153.1; JOINED.
EMBL; AF346710; AAK31153.1; JOINED.
EMBL; AF257182; AAF68989.1; -
HSP; Q9BZR6; 10ZN.
Genew; HGNC:13299; GPR48.
MIM; 606666; -
InterPro; IPR000276; GPCR_Rhodpsn.
InterPro; IPR002131; Ghrmn_receptor.
InterPro; IPR001611; LRR.
InterPro; IPR000372; LRR_Nterm.
InterPro; IPR003591; LRR_typ.
Pfam; PF00001; 7tm_1; 1.
Pfam; PF00560; LRR; 15.
Pfam; PF01462; LRRNT; 1.
PRINTS; PR00373; GLYCHORMONER.
PRINTS; PR00237; GPCR_RHODOPSN.
PRINTS; PR00019; LEURICHRPT.
SMART; SM00013; LRRNT; 1.
SMART; SM00369; LRR_TYP; 4.
PROSITE; PS00237; G_PROTEIN_REC_F1_1; FALSE_NEG.
PROSITE; PS0262; G_PROTEIN_REC_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
KW Signal; Transmembrane.
FT SIGNAL 1 24 Potential.
FT CHAIN 25 951 Leucine-rich repeat-containing G protein-
coupled receptor 4.
FT DOMAIN 25 544 Extracellular (Potential).
FT TRANSMEM 545 565 1 (Potential).
FT DOMAIN 566 575 Cytoplasmic (Potential).

```

```

FT TRANSMEM 576 596 2 (Potential).
FT DOMAIN 597 620 Extracellular (Potential).
FT TRANSMEM 621 641 3 (Potential).
FT DOMAIN 642 661 Cytoplasmic (Potential).
FT TRANSMEM 662 682 4 (Potential).
FT DOMAIN 683 703 Extracellular (Potential).
FT TRANSMEM 704 724 5 (Potential).
FT DOMAIN 725 756 Cytoplasmic (Potential).
FT TRANSMEM 757 777 6 (Potential).
FT DOMAIN 778 783 Extracellular (Potential).
FT TRANSMEM 784 804 7 (Potential).
FT DOMAIN 805 951 Cytoplasmic (Potential).
FT REPEAT 55 79 LRR 1.
FT REPEAT 81 103 LRR 2.
FT REPEAT 104 127 LRR 3.
FT REPEAT 128 151 LRR 4.
FT REPEAT 153 175 LRR 5.
FT REPEAT 176 199 LRR 6.
FT REPEAT 201 223 LRR 7.
FT REPEAT 224 247 LRR 8.
FT REPEAT 248 270 LRR 9.
FT REPEAT 272 294 LRR 10.
FT REPEAT 318 341 LRR 11.
FT REPEAT 342 366 LRR 12.
FT REPEAT 368 387 LRR 13.
FT REPEAT 388 411 LRR 14.
FT REPEAT 413 435 LRR 15.
FT DISULFID 618 693 By similarity.
FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 294 294 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 505 505 N-linked (GlcNAc...) (Potential).
FT CONFLICT 292 292 F -> S (in Ref. 1; AAF68989).
FT CONFLICT 433 433 L -> P (in Ref. 1; AAF68989).
FT CONFLICT 668 668 L -> S (in Ref. 1; AAF68989).
SQ SEQUENCE 951 AA; 104460 MW; 5E0C2DFCF22CA1BB CRC64;

Query Match 1.4%; Score 13; DB 1; Length 951;
Best Local Similarity 100.0%; Pred. No. 0.0078;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 NLSSLVVLHNN 220
|||||
DB 199 NLSSLVVLHNN 211
|||||

RESULT 12
LGR4_RAT STANDARD; PRT; 951 AA.
AC Q9Z2H4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
GN Name=Gpr48; Synonyms=Lgr4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;
RA Hsu S.Y., Liang S.-G., Hsueh A.J.W.;
RT "Characterization of two LGR genes homologous to gonadotropin and
thyrotropin receptors with extracellular leucine-rich repeats and a G
protein-coupled, seven-transmembrane region.";
RL Mol. Endocrinol. 12:1830-1845(1998).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.

```



```
FT  NON TER      1      1
SQ  NON TER     162     162
SQ  SEQUENCE   162 AA;  17608 MW;  C593128C551FB824 CRC64;

Query Match      1.2%; Score 11; DB 2; Length 162;
Best Local Similarity 100.0%; Pred.No. 0.15;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  810 PLPACLNPLLY 820
Db  70 PLPACLNPLLY 80

RESULT 15
Q8R301
ID  Q8R301      PRELIMINARY;      PRT;   459 AA.
AC  Q8R301;
DT  01-JUN-2002 (TReMBLrel. 21, Created)
DT  01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT  01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE  Lgr6 protein (Fragment).
GN  Name=Lgr6;
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Czech II; TISSUE=Mammary tumor;
RX  MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA  Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Raba S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA  Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA  Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA  Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
RA  Jones S.J., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length human
RL  and mouse cDNA sequences.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=Czech II; TISSUE=Mammary tumor;
RA  Strausberg R.;
RL  Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR  EMBL; BC026896; AAH26896.1; -.
DR  MGB; MGI:2441805; Lgr6.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR  GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR  InterPro; IPR000276; GPCR_Rhodopsn.
DR  InterPro; IPR002131; Gqhrna_receptor.
DR  Pfam; PF00001; 7cm1; 1.
DR  PRINTS; PR00373; GLYCHORMONER.
DR  PRINTS; PR00237; GPCRHHODPSN.
FT  NON TER      1
SQ  SEQUENCE   459 AA;  47889 MW;  F0100BF073E81762 CRC64;

Query Match      1.2%; Score 11; DB 2; Length 459;
Best Local Similarity 100.0%; Pred.No. 0.36;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  810 PLPACLNPLLY 820
```

Db 309 PLPACLNPLLY 319

|||||||

Search completed: July 12, 2005, 07:53:08
Job time : 114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:48:53 ; Search time 25 Seconds
(without alignments)
2708.268 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 907
Sequence: 1 MDTSLRGVLLSLPVLLQLAT.....PAYVTECHLSVAVPCL 907

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

- Database : Issued Patents AA:*
- 1: /cgn2_6/prodata/1/1aa/5A COMB.pep:*
 - 2: /cgn2_6/prodata/1/1aa/5B COMB.pep:*
 - 3: /cgn2_6/prodata/1/1aa/6A COMB.pep:*
 - 4: /cgn2_6/prodata/1/1aa/6B COMB.pep:*
 - 5: /cgn2_6/prodata/1/1aa/PCTUS COMB.pep:*
 - 6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	US-09-170-496D-264	Sequence 264, App
2	806	88.9	907	US-09-170-496D-278	Sequence 278, App
3	9	1.0	644	US-08-866-757-2	Sequence 2, Appli
4	9	1.0	644	US-09-153-593-2	Sequence 2, Appli
5	9	1.0	1133	US-10-101-464A-809	Sequence 809, App
6	8	0.9	99	US-09-482-273-258	Sequence 258, App
7	8	0.9	207	US-09-482-273-167	Sequence 167, App
8	8	0.9	232	US-09-543-681A-5153	Sequence 5153, App
9	8	0.9	259	US-09-482-273-260	Sequence 260, App
10	8	0.9	386	US-09-489-039A-10334	Sequence 10334, A
11	8	0.9	448	US-09-361-443-4	Sequence 4, Appli
12	8	0.9	467	US-09-361-443-2	Sequence 2, Appli
13	8	0.9	469	US-09-198-452A-793	Sequence 793, App
14	8	0.9	469	US-09-438-185A-745	Sequence 745, App
15	8	0.9	728	US-09-949-016-10391	Sequence 10391, A
16	7	0.8	28	US-09-099-041A-19	Sequence 19, Appl
17	7	0.8	28	US-09-245-281-19	Sequence 19, Appl
18	7	0.8	28	US-09-207-359B-19	Sequence 19, Appl
19	7	0.8	28	US-09-340-620A-19	Sequence 19, Appl
20	7	0.8	28	US-09-865-364-19	Sequence 19, Appl
21	7	0.8	36	US-08-216-592A-22	Sequence 22, Appl
22	7	0.8	67	US-09-248-796A-21331	Sequence 21331, A
23	7	0.8	69	US-09-910-009A-158	Sequence 158, App
24	7	0.8	88	US-09-248-796A-23648	Sequence 23648, A
25	7	0.8	95	US-09-328-352-5647	Sequence 5647, Ap
26	7	0.8	96	US-09-513-999C-4592	Sequence 4592, Ap
27	7	0.8	97	US-09-513-999C-7456	Sequence 7456, Ap

ALIGNMENTS

RESULT 1

US-09-170-496D-264
; Sequence 264, Application US/09170496D
; Patent No. 6555339

GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.

; APPLICANT: Chalmers, Derek T.

; APPLICANT: Liaw, Chen W.

; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein

; TITLE OF INVENTION: Receptors

; FILE REFERENCE: AREN-0040

; CURRENT APPLICATION NUMBER: US/09/170,496D

; CURRENT FILING DATE: 1998-10-13

; NUMBER OF SEQ ID NOS: 294

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 264

; LENGTH: 907

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-170-496D-264

Query Match		100.0%	Score 907;	DB 4;	Length 907;
Best Local Similarity		100.0%	Pred. No. 0;		
Matches	907;	Conservative	0;	Mismatches	0;
		Indels	0;	Gaps	0;
QY	1	MDTSRLGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL	60		
Db	1	MDTSRLGVLLSLPVLLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL	60		
QY	61	PNLSVFTSYLDLSNNITSQLPNPLSLRFLFEELRAGNALTYIPKGAFTGLYSIKVLM	120		
Db	61	PNLSVFTSYLDLSNNITSQLPNPLSLRFLFEELRAGNALTYIPKGAFTGLYSIKVLM	120		
QY	121	LQNNQLRHVPTALQNLRSIQSLRDANHI SVVPSCFGLSHRLWLDNLTALTEIPVQ	180		
Db	121	LQNNQLRHVPTALQNLRSIQSLRDANHI SVVPSCFGLSHRLWLDNLTALTEIPVQ	180		
QY	181	AFRSLAQAMTALANKIHHPDYAFGNLSLVVLHLHNNRHS LGKKCFDGLHSLD	240		
Db	181	AFRSLAQAMTALANKIHHPDYAFGNLSLVVLHLHNNRHS LGKKCFDGLHSLD	240		
QY	241	LYNNILDEPPTAIRLSNLKELGFHSNNIRSIPEKAFVGNPNSLIITHFDNPIQFVGRSA	300		
Db	241	LYNNILDEPPTAIRLSNLKELGFHSNNIRSIPEKAFVGNPNSLIITHFDNPIQFVGRSA	300		
QY	301	FOHLPRLTLTUNGASQITEFPDLTGNTANLESITLTGAQISSLPQVCNQLPNQLVDLS	360		
Db	301	FOHLPRLTLTUNGASQITEFPDLTGNTANLESITLTGAQISSLPQVCNQLPNQLVDLS	360		
QY	361	YNLLEDPSFSVQCKLQKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIAIHPNAFST	420		

28	7	0.8	113	4	US-08-816-772-2	Sequence 2, Appli
29	7	0.8	115	4	US-09-640-211A-2235	Sequence 2235, Ap
30	7	0.8	127	3	US-09-134-001C-4041	Sequence 4041, Ap
31	7	0.8	135	4	US-10-101-464A-563	Sequence 563, App
32	7	0.8	137	4	US-09-489-039A-11223	Sequence 11223, A
33	7	0.8	149	4	US-09-902-540-11269	Sequence 11269, A
34	7	0.8	154	4	US-09-540-236-2694	Sequence 2694, Ap
35	7	0.8	156	4	US-09-252-991A-23000	Sequence 23000, A
36	7	0.8	157	4	US-09-270-767-40195	Sequence 40195, A
37	7	0.8	157	4	US-09-270-767-55411	Sequence 55411, A
38	7	0.8	161	4	US-09-248-796A-15595	Sequence 15595, A
39	7	0.8	161	4	US-10-101-464A-570	Sequence 570, App
40	7	0.8	166	4	US-09-252-991A-21079	Sequence 21079, A
41	7	0.8	166	4	US-09-889-914B-8	Sequence 8, Appli
42	7	0.8	168	4	US-09-270-767-57047	Sequence 57047, A
43	7	0.8	170	3	US-08-875-573-6	Sequence 6, Appli
44	7	0.8	176	4	US-09-134-000C-4892	Sequence 4892, Ap
45	7	0.8	179	4	US-09-252-991A-24661	Sequence 24661, A

Db 361 YNLLDLPFSVCQKQKIDLRHNEIYEIKVDTFQQLSLSLNLANWKIAIIHNAFST 420
Qy 421 LPSLTKLSSNLLSFPITGLHGLTHLKLTCGNHALQSLISSENPELKVIEPYAYQCC 480
Db 421 LPSLTKLSSNLLSFPITGLHGLTHLKLTCGNHALQSLISSENPELKVIEPYAYQCC 480
Qy 481 AFGVCENAYKISQNWKNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCEHLLDGWLIIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCEHLLDGWLIIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGVSSAVLAGVDAFTFGSFARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSSAVLAGVDAFTFGSFARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSAKFETKAPFSSLKVIIICALLATWAAVPLLGGSYKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLKVIIICALLATWAAVPLLGGSYKYGASPLCLPLPFGEP 720
Qy 721 STMGYMWALILLNSLCFLMMTTIATKLYCNLDKGDLENIDWCSWKHIALLLFTNCILNC 780
Db 721 STMGYMWALILLNSLCFLMMTTIATKLYCNLDKGDLENIDWCSWKHIALLLFTNCILNC 780
Qy 781 PVAFLSFSLLNLTFISPEVVKFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Db 781 PVAFLSFSLLNLTFISPEVVKFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDVKKQSCDSTQALVFTTSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVKKQSCDSTQALVFTTSSITVDLPSSVPSPAYPVTESCHLSS 900
Qy 901 VAFVPCCL 907
Db 901 VAFVPCCL 907

RESULT 2
US-09-170-496D-278
; Sequence 278 Application US/09170496D
; Patent No. 655339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 655339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 278
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-278

Query Match 88.9%; Score 806; DB 4; Length 907;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 906; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MDTSRIGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Db 1 MDTSRIGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSEL 60
Qy 61 PSNLSVFTSYLDLSMNINSQLLPNPLPSLRFLEELRLAGNALTYIPKGFTGLYSKLVLM 120
Db 61 PSNLSVFTSYLDLSMNINSQLLPNPLPSLRFLEELRLAGNALTYIPKGFTGLYSKLVLM 120

Qy 121 LONNOLRHVPTEALQNLRSLOSLRLDANHISYVPPSCFSGHLHSLRHLWLDNALTEIPVQ 180
Db 121 LONNOLRHVPTEALQNLRSLOSLRLDANHISYVPPSCFSGHLHSLRHLWLDNALTEIPVQ 180
Qy 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFDGLHSLD 240
Db 181 AFRSLSALQAMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFDGLHSLD 240
Qy 241 LNYNNLDFFPTAIRTLNKLGLGPHSNNIRSIPEKAFVGNFSLTIHFYDNPFIQVGRSA 300
Db 241 LNYNNLDFFPTAIRTLNKLGLGPHSNNIRSIPEKAFVGNFSLTIHFYDNPFIQVGRSA 300
Qy 301 FOHLPELRTLTNGASQITTEPDLTGTANLSLTLTGAQISSLPOTVCNQLPNLQVLDLS 360
Db 301 FOHLPELRTLTNGASQITTEPDLTGTANLSLTLTGAQISSLPOTVCNQLPNLQVLDLS 360
Qy 361 YNLLDLPFSVCQKQKIDLRHNEIYEIKVDTFQQLSLSLNLANWKIAIIHNAFST 420
Db 361 YNLLDLPFSVCQKQKIDLRHNEIYEIKVDTFQQLSLSLNLANWKIAIIHNAFST 420
Qy 421 LPSLTKLSSNLLSFPITGLHGLTHLKLTCGNHALQSLISSENPELKVIEPYAYQCC 480
Db 421 LPSLTKLSSNLLSFPITGLHGLTHLKLTCGNHALQSLISSENPELKVIEPYAYQCC 480
Qy 481 AFGVCENAYKISQNWKNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ 540
Db 481 AFGVCENAYKISQNWKNKGDNSMDDLHKDAGMFOAQDERDLEDFLLDPEEDLKALHSVQ 540
Qy 541 CSPSPGPKPCEHLLDGWLIIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCEHLLDGWLIIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGVSSAVLAGVDAFTFGSFARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
Db 601 AVNMLTGVSSAVLAGVDAFTFGSFARHGAWENGVGCHVIGFLSIFASESSVFLTLAAL 660
Qy 661 ERGFSVKYSAKFETKAPFSSLKVIIICALLATWAAVPLLGGSYKYGASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAKFETKAPFSSLKVIIICALLATWAAVPLLGGSYKYGASPLCLPLPFGEP 720
Qy 721 STMGYMWALILLNSLCFLMMTTIATKLYCNLDKGDLENIDWCSWKHIALLLFTNCILNC 780
Db 721 STMGYMWALILLNSLCFLMMTTIATKLYCNLDKGDLENIDWCSWKHIALLLFTNCILNC 780
Qy 781 PVAFLSFSLLNLTFISPEVVKFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Db 781 PVAFLSFSLLNLTFISPEVVKFILLVVVPLPACLNPLLYILFNPHEKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDVKKQSCDSTQALVFTTSSITVDLPSSVPSPAYPVTESCHLSS 900
Db 841 WTRSKHPSLMSINSDVKKQSCDSTQALVFTTSSITVDLPSSVPSPAYPVTESCHLSS 900
Qy 901 VAFVPCCL 907
Db 901 VAFVPCCL 907
RESULT 3
US-08-866-757-2
; Sequence 2, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGMA, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA


```
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-866-757-2

Query Match 1.0%; Score 9; DB 2; Length 644;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 474 PYAYQCCAF 482
Db 158 PYAYQCCAF 166

RESULT 4
US-09-153-593-2
; Sequence 2, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOUBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSA, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CNA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 644
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; US-09-153-593-2

Query Match 1.0%; Score 9; DB 3; Length 644;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 474 PYAYQCCAF 482
Db 158 PYAYQCCAF 166

RESULT 5
US-10-101-464A-809
; Sequence 809, Application US/10101464A

; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 809
; LENGTH: 1133
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
; US-10-101-464A-809

Query Match 1.0%; Score 9; DB 4; Length 1133;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 LQVLDLSYN 362
Db 175 LQVLDLSYN 183

RESULT 6
US-09-482-273-258
; Sequence 258, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-482-273-258

Query Match 0.9%; Score 8; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 YTKLYCNL 751
Db 33 YTKLYCNL 40

RESULT 7
US-09-482-273-167
; Sequence 167, Application US/09482273
```


; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 167
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-167

Query Match 0.9%; Score 8; DB 4; Length 207;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 810 PLPACLPN 817
|||||||
DB 47. PLPACLPN 54

RESULT 8

US-09-543-681A-5153
; Sequence 5153, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5153
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5153

Query Match 0.9%; Score 8; DB 4; Length 232;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 FTSYLDLS 74
|||||||
DB 220 FTSYLDLS 227

RESULT 9

US-09-482-273-260
; Sequence 260, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: PZ030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14

; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 260
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-260

Query Match 0.9%; Score 8; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 YTKLYCNL 751
|||||||
DB 33 YTKLYCNL 40

RESULT 10

US-09-489-039A-10334
; Sequence 10334, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10334
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10334

Query Match 0.9%; Score 8; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 830 DLVSLRKQ 837
|||||||
DB 352 DLVSLRKQ 359

RESULT 11

US-09-361-443-4
; Sequence 4, Application US/09361443
; Patent No. 6660275
; GENERAL INFORMATION:
; APPLICANT: Murdin, Andrew D
; APPLICANT: Omen, Raymond P
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 19721-005 (AV-5)
; CURRENT APPLICATION NUMBER: US/09/361,443
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/094,195
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 448
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-361-443-4


```
Query Match          0.9%; Score 8; DB 4; Length 448;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 SFSSLINL 793
DB 290 SFSSLINL 297

RESULT 12
US-09-361-443-2
; Sequence 2, Application US/09361443
; Patent No. 6660275
; GENERAL INFORMATION:
; APPLICANT: Murdin, Andrew D
; APPLICANT: Oomen, Raymond P
; TITLE OF INVENTION: CHLAMYDIA ANTIGENS AND CORRESPONDING DNA FRAGMENTS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 19721-005 (AV-5)
; CURRENT APPLICATION NUMBER: US/09/361,443
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/094,195
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 467
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-361-443-2

Query Match          0.9%; Score 8; DB 4; Length 467;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 SFSSLINL 793
DB 292 SFSSLINL 299

RESULT 13
US-09-198-452A-793
; Sequence 793, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifffais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; TITLE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 793
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-793

Query Match          0.9%; Score 8; DB 4; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 SFSSLINL 793
DB 294 SFSSLINL 301

RESULT 14
US-09-438-185A-745
; Sequence 745, Application US/09438185A
; Patent No. 6822071
```

```
; GENERAL INFORMATION:
; APPLICANT: Stephens, Richard
; APPLICANT: Mitchell, Wayne
; APPLICANT: Kalman, Sue
; APPLICANT: Davis, Ronald
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Chlamydia Pneumoniae Genome Sequence
; FILE REFERENCE: 018941-00041105
; CURRENT APPLICATION NUMBER: US/09/438,185A
; CURRENT FILING DATE: 2002-03-13
; PRIOR APPLICATION NUMBER: US 60/108,279
; PRIOR FILING DATE: 1998-11-12
; PRIOR APPLICATION NUMBER: US 60/128,606
; PRIOR FILING DATE: 1999-04-08
; NUMBER OF SEQ ID NOS: 1074
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 745
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
; FEATURE:
; OTHER INFORMATION: CPn0743
US-09-438-185A-745

Query Match          0.9%; Score 8; DB 4; Length 469;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 SFSSLINL 793
DB 294 SFSSLINL 301

RESULT 15
US-09-949-016-10391
; Sequence 10391, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10391
; LENGTH: 728
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10391

Query Match          0.9%; Score 8; DB 4; Length 728;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 650 SSVFLLTL 657
DB 580 SSVFLLTL 587

Search completed: July 12, 2005, 07:54:11
Job time : 28 secs
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:30 ; Search time 31 Seconds
(without alignments)
2815.116 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSLRGVLLSLPVLQLAT.....PAYPTESCHLSSVAFVCL 907

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4702	100.0	907	2 JG0176	orphan G protein-c
2	4027	85.6	907	2 JG0193	G protein-coupled
3	650	13.8	695	2 I45896	follicle stimulat
4	648.5	13.8	695	1 JN0898	follicle stimulat
5	642.5	13.7	694	2 JC2237	follicle stimulat
6	642	13.7	695	2 JC1493	follicle stimulat
7	636	13.5	692	2 A34548	follicle stimulat
8	632	13.4	694	2 JC4301	follicle stimulat
9	625	13.3	696	2 JC7361	follicle stimulat
10	621.5	13.2	695	1 QRHUFT	follicle stimulat
11	600	12.8	700	2 A42395	lutropin receptor
12	594	12.6	764	1 QRHURH	thyrotropin recept
13	591	12.6	764	2 A40077	thyrotropin recept
14	588.5	12.5	700	2 I77463	luteinizing hormon
15	588.5	12.5	700	2 A49744	lutropin-choriogon
16	581	12.4	764	2 I48882	thyrotropin recept
17	578	12.3	793	2 JC7390	thyroid stimulat
18	572	12.2	696	2 A41344	lutropin-choriogon
19	567.5	12.1	764	2 JC5643	thyroid stimulat
20	566.5	12.0	764	2 A35956	thyrotropin recept
21	560	11.9	699	1 QRHUFT	lutropin-choriogon
22	536.5	11.4	814	2 JC7389	thyroid stimulat
23	508.5	10.8	889	2 T20123	hypothetical prote
24	490	10.4	925	2 JC2033	G protein-coupled
25	426.5	9.1	603	2 JC1282	insulin-like growt
26	421	9.0	605	2 JC5239	insulin-like growt
27	418.5	8.9	622	2 JC7973	synleucin - human
28	406.5	8.6	603	2 JC6128	insulin-like growt
29	406.5	8.6	1389	2 TJ3852	gene wheeler prote

30	403	8.6	1523	2 T13953	MEGF5 protein - ra
31	402	8.5	605	2 A41915	insulin-like growt
32	400.5	8.5	1385	2 T13887	tlr protein - frui
33	398	8.5	1091	2 A58532	glial cell membran
34	391.5	8.3	1469	2 B36665	slit protein 2 pre
35	391.5	8.3	1480	2 A36665	slit protein 1 pre
36	385.5	8.2	1531	2 T42218	slit-1 protein hom
37	382.5	8.1	1066	2 T15864	hypothetical prote
38	381.5	8.1	1115	2 S40241	G protein-coupled
39	371	7.9	1119	2 AD1822	leucine-rich-repea
40	370	7.9	559	2 T42998	ras-binding protei
41	370	7.9	560	2 A60164	platelet membrane
42	364.5	7.8	572	2 T30947	hypothetical prote
43	360.5	7.7	1134	1 A29944	chaoptin precursor
44	356.5	7.6	536	2 A34901	lysine carboxypept
45	354.5	7.5	707	2 JC7763	neuronal leucine-r

ALIGNMENTS

RESULT 1

JE0176
orphan G protein-coupled receptor precursor - human
C:Species: Homo sapiens (man)
C:Date: 03-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 09-Jul-2004
C:Accession: JE0176
R:McDonald, T.; Wang, R.; Bailey, W.; Xie, G.; Chen, F.; Caskey, C.T.; Liu, Q.
Biochem. Biophys. Res. Commun. 247, 266-270, 1998
A:Title: Identification and cloning of an orphan G protein-coupled receptor of the glyco
A:Reference number: JE0176; MUID:98308104; PMID:9642114
A:Accession: JE0176
A:Molecule type: mRNA
A:Residues: 1-907 <MCD>
A:Cross-references: UNIPROT:O75473; GB:AF062006; NID:G3366801; PIDN:AAC28019.1; PID:G33
C:Comment: This protein is a receptor for a novel class of glycoprotein ligands.
C:Genetics:
A:Gene: HG38
A:Map position: 12Q22-23
F:1-21/Domain: signal sequence #status predicted <SIG>
F:562-583/Domain: transmembrane #status predicted <TM1>
F:594-616/Domain: transmembrane #status predicted <TM2>
F:639-660/Domain: transmembrane #status predicted <TM3>
F:681-701/Domain: transmembrane #status predicted <TM4>
F:725-744/Domain: transmembrane #status predicted <TM5>
F:768-791/Domain: transmembrane #status predicted <TM6>
F:803-824/Domain: transmembrane #status predicted <TM7>

Query Match	100.0%	Score	4702	DB	2	Length	907
Best Local Similarity	100.0%	Pred. No.	4.6e-297				
Matches	907	Conservative	0	Mismatches	0	Indels	0
Gaps	0						
QY	1	MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL	60				
Db	1	MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDGLSEL	60				
QY	61	PSNLSVFTSYLDLSNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKVLVM	120				
Db	61	PSNLSVFTSYLDLSNNISQLLPNPLPSLRFLEELRAGNALTYPKGAFTGLYSKVLVM	120				
QY	121	LQNNQLRHVPTEALQNLSLQSLRDANHISVPPSCFSGLSRLHRLWDDNALTEIPVQ	180				
Db	121	LQNNQLRHVPTEALQNLSLQSLRDANHISVPPSCFSGLSRLHRLWDDNALTEIPVQ	180				
QY	181	AFRSLSALQMTALANKIHHIPDYAFGNLSLVLLHNNRHISLGKKCFDGLHSLTLD	240				
Db	181	AFRSLSALQMTALANKIHHIPDYAFGNLSLVLLHNNRHISLGKKCFDGLHSLTLD	240				
QY	241	LYNNNLDEFPTAIRTLNKLGFHNNIRISIRPEKAFVGNPSLTIHFVNDPIQVGRSA	300				
Db	241	LYNNNLDEFPTAIRTLNKLGFHNNIRISIRPEKAFVGNPSLTIHFVNDPIQVGRSA	300				
QY	301	FQHLPELRTLINGASQITEFPDLTGTANLESLLTGAQISSLPQVCNQLPNLQVLDLS	360				


```
Db 301 FOHLPELRTLTNGASQITPEPDLTGTAHNSLTLTGAISSLQTVTCNQLPNQVLDLS 360
Qy 361 YNLEDLPSFVSCQKLOKIDLRHNEIYEIKVDYTFQQLSLSLSLNLANWKIAIHPNAPST 420
Db 361 YNLEDLPSFVSCQKLOKIDLRHNEIYEIKVDYTFQQLSLSLSLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKDLSSNLSSFPITGLHGLTHLKTGNHALQSLISSENPELKVTEMVAYOCC 480
Db 421 LPSLIKDLSSNLSSFPITGLHGLTHLKTGNHALQSLISSENPELKVTEMVAYOCC 480
Qy 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Qy 541 CSPSPGPKPCEHLLDGLWLRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCEHLLDGLWLRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGSSAVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLTAAAL 660
Db 601 AVNMLTGSSAVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLTAAAL 660
Qy 661 ERGFSVKYSAXFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAXFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Qy 721 STMGYVALIILNSLCFLMMTIAVTKLYCNDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNSLCFLMMTIAVTKLYCNDKGDLENWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSPSSNLNLTFTISPEVVKFILLVVPVLPACLNPLLYLTFNPHKEDLVSLRKQTYV 840
Db 781 PVAFLSPSSNLNLTFTISPEVVKFILLVVPVLPACLNPLLYLTFNPHKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907

RESULT 2
JG0193
G protein-coupled receptor FEX - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JG0193
R:Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A:Title: Identification of a novel seven-transmembrane receptor with homology to glycophorin
A:Reference number: JG0193; MUID:99121227; PMID:9920770
A:Accession: JG0193
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HER>

Query Match 85.6%; Score 4027; DB 2; Length 907;
Best Local Similarity 85.3%; Pred No. 2, 8e-253;
Matches 774; Conservative 53; Mismatches 80; Indels 0; Gaps 0;

Qy 1 MDTSLRGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRLMRVDCSDLGISEL 60
Db 1 MDTSCVHMLLSLALLQLVAAGSSPGDPAIPRGCPCHCELDGRLMRVDCSDLGISEL 60
Qy 61 PSNLSVFTSYDLSNMNNTSOLLPNPLPSLFLLELRAGNALTYPKGAFTGLYSKVLVM 120
Db 61 PSNLSVFTSYDLSNMNNTSOLLPNPLPSLFLLELRAGNALTYPKGAFTGLYSKVLVM 120
Qy 121 LQNNOLRVHPTEALQNLRSLSQLRDANHISVYPPSCFSGHLSHRLWLDNALTDPVQ 180
Db 121 LQNNOLRVHPTEALQNLRSLSQLRDANHISVYPPSCFSGHLSHRLWLDNALTDPVQ 180
```

```
Db 121 LQNNOLRVHPTEALQNLRSLSQLRDANHISVYPPSCFSGHLSHRLWLDNALTDPVQ 180
Qy 181 AFRSLSALQMTALANKIHHIPDYAFGNLSLVLVHLHNNRIHSLGKKCFDGLHSLETLD 240
Db 181 AFRSLSALQMTALANKIHHIPDYAFGNLSLVLVHLHNNRIHSLGKKCFDGLHSLETLD 240
Qy 241 LAYNNLDEFPFTAIRTLNKLKELGPHSNNIRSIPEKAFVGNPSLTIITIFYDNPIQFVGRSA 300
Db 241 LAYNNLDEFPFTAIRTLNKLKELGPHSNNIRSIPEKAFVGNPSLTIITIFYDNPIQFVGRSA 300
Qy 301 FOHLPELRTLTNGASQITPEPDLTGTAHNSLTLTGAISSLQTVTCNQLPNQVLDLS 360
Db 301 FOHLPELRTLTNGASQITPEPDLTGTAHNSLTLTGAISSLQTVTCNQLPNQVLDLS 360
Qy 361 YNLEDLPSFVSCQKLOKIDLRHNEIYEIKVDYTFQQLSLSLSLNLANWKIAIHPNAPST 420
Db 361 YNLEDLPSFVSCQKLOKIDLRHNEIYEIKVDYTFQQLSLSLSLNLANWKIAIHPNAPST 420
Qy 421 LPSLIKDLSSNLSSFPITGLHGLTHLKTGNHALQSLISSENPELKVTEMVAYOCC 480
Db 421 LPSLIKDLSSNLSSFPITGLHGLTHLKTGNHALQSLISSENPELKVTEMVAYOCC 480
Qy 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Db 481 AFGVCENAYKISNOWNKGDNSMDLHKDAGMFOAQDERDLEFLDFFEDLKALHSVQ 540
Qy 541 CSPSPGPKPCEHLLDGLWLRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Db 541 CSPSPGPKPCEHLLDGLWLRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIA 600
Qy 601 AVNMLTGSSAVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLTAAAL 660
Db 601 AVNMLTGSSAVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLTAAAL 660
Qy 661 ERGFSVKYSAXFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Db 661 ERGFSVKYSAXFETKAPFSSLKVIILLCALLATMAAVPLLGSKYKASPLCLPLPFGEP 720
Qy 721 STMGYVALIILNSLCFLMMTIAVTKLYCNDKGDLENWDCSMVKHIALLLFTNCILNC 780
Db 721 STMGYVALIILNSLCFLMMTIAVTKLYCNDKGDLENWDCSMVKHIALLLFTNCILNC 780
Qy 781 PVAFLSPSSNLNLTFTISPEVVKFILLVVPVLPACLNPLLYLTFNPHKEDLVSLRKQTYV 840
Db 781 PVAFLSPSSNLNLTFTISPEVVKFILLVVPVLPACLNPLLYLTFNPHKEDLVSLRKQTYV 840
Qy 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Db 841 WTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITVDLPSSVPSPAYPVPTESCHLSS 900
Qy 901 VAFVPCPL 907
Db 901 VAFVPCPL 907

RESULT 3
I45896
follicle stimulating hormone receptor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
C:Accession: I45896
R:Houde, A.; Lambert, A.; Saumande, J.; Silversides, D.W.; Lussier, J.G.
Mol. Reprod. Dev. 39, 127-135, 1994
A:Title: Structure of the bovine follicle-stimulating hormone receptor complementary DNA
A:Reference number: I45896; MUID:95127199; PMID:7826612
A:Accession: I45896
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-895 <HOU>
C:Cross-references: UNIPROT:P35376; GB:L22319; NID:g404671; PIDN:AAC37324.1; PID:g404672
C:Genetics:
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
```


A;Cross-references: UNIPROT:P35379
A;Experimental source: testis
R;Yarney, T.A.; Sairam, M.R.; Khan, H.; Ravindranath, N.; Payne, S.; Seidah, N.G.
Mol. Cell. Endocrinol. 93, 219-226, 1993
A;Title: Molecular cloning and expression of the ovine testicular follicle stimulating h
A;Reference number: I47080; MUID:93351750; PMID:8394255
A;Accession: I47080
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-695 <YAR>
A;Cross-references: GB:L07302; NID:g165884; PIDN:AAA31525.1; PID:g165885
C;Genetics:
A;Gene: PSH-R
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F;191,199/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.7%; Score 642; DB 2; Length 695;
Best Local Similarity 23.0%; Pred. No. 8.le-34;
Matches 213; Conservative 139; Mismatches 304; Indels 270; Gaps 27;

Qy 10 LSLPVLQLATGSSPSRSGVLLRGCPH-CHCEPDGRLRLRVDCDLGLSELPSNLSVFT 68
Db 3 LFLVALLAFSLGS-----GCHRLCHCS-NGVFL-----CDQSKVTMPSD----- 43
Qy 69 SYLDLSMNNISQLLPNPLPSLRP-LEELRLAGNALTYIPKGAFTGLYSL-KVLMQLNNOL 126
Db 44 -----LPRDAVELRFVLTCLR-----IPEGAFSGFGDLKIEISQNDVL 83
Qy 127 RHVPTALONLSQSLRLD-ANHSYVPPSCFGLSHSLRHLDNLTALTEIPVQAFSL 185
Db 84 EVIEANVSNLPKLEIRIEKANNLYIDPPAFQNLRLVLLISNTGIKHLP--AVHKI 141
Qy 186 SALQAMTALANKIHHIPDYAFGNLSLVVHLHNNRIHSLGKCKPDGLHSLLETLDLNNY 245
Db 142 QSLQKVLDDIQ-----DNNIHTVENSFMGL----- 168
Qy 246 LDEFTAIRTLNLSKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPITQFVGRSAFOHLP 305
Db 169 --SPESMIVWLS-----KNGIQEIHCNFG----- 192
Qy 306 ELRTLTNGASQITFPDLTGTAANLESITLTGAQISSLPQVNCNQLPMLQVLDLSYNLLE 365
Db 193 -----TQDEL-NLSDNSNLE-----BLPN----- 211
Qy 366 DLPSFVCQKQKIDRLNEIYEIKVDTFQQLLSRLSLNLANWKTAITHNAPSTLPSLI 425
Db 212 -----DVFGAAGPV 221
Qy 426 KLDLSNLLSPFITGLHGLTHLKLGTGNHALQSLISSENFPELKVIEPMFYAQCAFVGC 485
Db 222 ILDISRTIRSLPSVGLNKKLRKASYHLKLPSEKFEVTLVEASLTYPHCCAF-- 279
Qy 486 ENAYKISNQWNGDNSSMDLH-----KQAGMFOQDER-----DLE 523
Db 280 -----NWR-----QTSDLHPICNKSILRQEVDDMTQARQISLAEDDEPSYAKGFD 327
Qy 524 DFLDPEEDL-KALHSVCSPPGPKCEHLDDGLIRIGVWTIATVLATCNALVTSTV 582
Db 328 MMYSEFDYDLCEVVDVTCSPEDAFNCPEDIMGVDIRLVITFISILATGNILVLVIL 387
Qy 583 FRSPLYIPKILLIGVIAVNMILGVSAVLAVGDAFTFGSFARCAWNGVGVCHVIGF 642
Db 388 ITSQKLTVPRLMCLNLAFLADLCIGIYLLLASVDVHTKSYHNYAIDWQTGAGCDAGF 447
Qy 643 LSIFASESVFLTLTAALRGFSVKYSAKFETKAPFSKVLVILLCALLATMAAVPLLG 702
Db 448 FTVFASLSVYTLTATILTRWHITATHMQLECKVHVRHAASIMLVGVWFVFAVALFPFIFG 507
Qy 703 GSKYGASPLCLPLPFGESTWGYVVALILLNSLCLPMWTIATYKLYCNLDKGLD-ENWD 761
Db 508 ISSYMKVSIPLPMDIDSLQSLYVMSLLVNLAFVWVIGCYTHIYLVTRNPNTSSSD 567

Qy 762 CSMVGHIALLTNCLNCPVAFSLFSSLLNLTFTSPDEVIKFILLVVVPLPACLNPLLYI 821
Db 568 TKIARMAWLIPTDFLCMAPISFAISASLKVPLITVSKSKILLVLFYPIINSCANPLFYA 627
Qy 822 LFNPFHKEDLVSLRKQTYVWTRSKHPSLMSINSDDDVEKOSCDSTQALVTFTSSSIT-YDL 880
Db 628 IFTNFRFRDFFIL-----LSKFGCYEVOAQYTRSETSTFAHNF 665
Qy 881 PPSSVPSPAYP-VTESCHLSSVAFVP 905
Db 666 HPRNGHCPPAPRVING---SNVTLP 688

RESULT 7

A34548
follicotropin receptor precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004
C;Accession: A34548; A41729
R;Sprengel, R.; Braun, T.; Nikolic, K.; Segaloff, D.L.; Seeburg, P.H.
Mol. Endocrinol. 4, 525-530, 1990
A;Title: The testicular receptor for follicle stimulating hormone: structure and function
A;Reference number: A34548; MUID:91125358; PMID:2126341
A;Accession: A34548
A;Molecule type: mRNA
A;Residues: 1-692 <SPR>
A;Cross-references: UNIPROT:P20395; GB:L02842; NID:g204183; PIDN:AAA41175.1; PID:g204184
R;Heckert, L.L.; Daley, I.J.; Griswold, M.D.
Mol. Endocrinol. 6, 70-80, 1992
A;Title: Structural organization of the follicle-stimulating hormone receptor gene.
A;Reference number: A41729; MUID:92149579; PMID:1738373
A;Accession: A41729
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-692 <HEC>
A;Cross-references: GB:S81199; NID:g245344; PIDN:AAB21415.1; PID:g245345
A;Note: sequence inconsistent with the nucleotide translation
R;Davis, D.; Liu, X.; Segaloff, D.L.
Mol. Endocrinol. 9, 159-170, 1995
A;Title: Identification of the sites of N-linked glycosylation on the follicle-stimulating hormone receptor
A;Reference number: A57562; MUID:95295729; PMID:7776966
A;Contents: annotation; glycosylation sites
C;Function:
A;Description: receptor that mediates the biochemical effects of follicotropin
C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece

F;16-692/Product: follictropin receptor #status predicted <NAT>
F;16-366/Domain: extracellular hormone binding #status predicted <EB>
F;56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;367-387/Domain: transmembrane #status predicted <TM1>
F;398-421/Domain: transmembrane #status predicted <TM2>
F;443-465/Domain: transmembrane #status predicted <TM3>
F;486-508/Domain: transmembrane #status predicted <TM4>
F;529-550/Domain: transmembrane #status predicted <TM5>
F;574-597/Domain: transmembrane #status predicted <TM6>
F;609-630/Domain: transmembrane #status predicted <TM7>
F;191,199,233/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;554/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;595/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match

13.5%; Score 636; DB 2; Length 692;

Best Local Similarity 23.1%; Pred. No. 2e-33;

Matches 211; Conservative 129; Mismatches 310; Indels 264; Gaps 25;

Qy	9	LLSLPVLQLQATGSSPSRGVLLRGCPTH---CHCEPDGRMLLRVDCSDLGJSELPNSLUS	65
Db	3	LLLVSLLAFLAGTGS-----GC--HHWLCHC--SNRVFL---CQDSKTEIPT---	42
Qy	66	VFTSVLDLSMNNISQLLPNPLPSLRFLERLRAGNALTVIPKGAFTGLYSL-KVLMLQNN	124
Db	43	-----DLPRNAI-----ELRFVLTKLRVIPKGSFAGFDGJKEIISQND	81
Qy	125	QLRHVPTTEALQNLRSLOSLRD-ANHISYVPPSCFSGLHSLRHLMDLDNALTEIPVQAFR	183
Db	82	VLEIVEADVFNLPKLEIRIEKANLLYINPEAFQNLPSRLAYLLISNTGIKHLP--AVH	139
Qy	184	SLSALOQMTALANKTHIHDPYAFGNLSLVVLHLHNNRIHSGKCKCFDGLHSLETLDLNY	243
Db	140	KIOSLQKVLLDIQ-----DNNITHIVARNSFMGL-----	168
Qy	244	NNLDEFPPTAIRTLNKLKELGFHSHNNIRSIPEKAFVGNPSLIITHFDYDNP IQFVGRSAFOH	303
Db	169	-----SFESVILWLVS-----XNGIEEIHNCAFNG-----	192
Qy	304	LPELRTLTLNGASQITEFPDLTGTALESRLTLTGAQISSLPTQVCNQLPNLQVLDLSYNL	363
Db	193	-----TQDELNLSD--NNN	205
Qy	364	LEDLPSPFVCOQKQIKDIRHNEIYEIKVDTPQOQLLSRLSLNLAWNKIAIHPNASTLPS	423
Db	206	LEELPN-----DVFGQASG	219
Qy	424	LIKDLSNLLSSPPIGLHGHTLKLTCNHALQSLISSENPELKVIMPEYAYOCCARG	483
Db	220	PVILDISRTKVHSLPNHGLENLKKURASTYRELKKPLNLDKFTVLMEASLTTPSHCCAP	279
Qy	484	-----VCENA---YKISNOWNKGDN--SSMDD--LHKDQAGMFQAOERDLEDF	525
Db	280	NLKRQISELHPICNKSILRQDIDDMTQIGDQRVSLIDDEPSYKSGSDMMYNFEYDLCNE	339
Qy	526	LLDPEEDLKALHSVOCSPSPGPKCEHLLDGWLIRIGVWTIAVLATCNALVTSTVFRS	585
Db	340	VVD-----VTCSPKPAFNPECEDIMGYNILRVLWIFISILAITGNTTTLVVVLTTSS	389
Qy	586	PLYISPIKLLIGVIAAVNMLTCVSSAVLAGVDAPFGSPARHGAWENGVCGVHGFLSI	645
Db	390	QYKLTVPFLMCLNLAFCALCGIYLLLASVDIHTKSOYHNIAIDMTQAGCDAAGFFTV	449
Qy	646	PASESSVPELLTALAALERGFSSVKYSAKFETKAPFSSKLWIIILLCALLATMAAVPLLGSK	705
Db	450	FASELSVYTLTAITLERWHTITHAMQLECKVLRHAASVWVLGWTFAPAAALFPFIGISS	509
Qy	706	YGASPLCLPLPFGPEBPTSGWYVVALILNLSLCFLMTIAVTKLYCNLDKGLD-ENIWDGSM	764
Db	510	YMKVSIICLPMBIDISPLSQLYVMALVNLVLAFFVVICGVCTHYIYLTVRNPFTIVSSSDTKI	569
Qy	765	VKHALLAFTNCILNCPVAELFSFSSLINLFTISPEVIFKILLVVVPLPACLNPPLIYILFN	824
Db	570	AKRWATLITFDPLCMAPISFPAISASLKVPLITVSKAKILLVFLTPINSKANPFLYAITF	629
Qy	825	PHFKEDLVSL-----RKQTVVWTRSRKHPSLMSINSDSDEKQSCDSTQOALT	870
Db	630	KNFRDRFDILLKFCGYEMQAOIYETETSSATHNFH-----ARKSHCSSAPRV---	677
Qy	871	FTSSSITVDLPPSS	884
Db	678	-TNSYVLVPLNHSS	690

RESULT 8

RESULT
JC4301

JC4301 follicleotropin receptor - nia

follicle-stimulating hormone receptor - pig

N;Alternate names: 1
C:Species: 5118 scrot

C;Species: Sus scrofa domestica (domestic pig)
C;Date: 16-Nov-1995 #sequence revision 08-Feb-1996 #text change 21-Jan-2000

C;Date: 16-Nov-1995
C;Accession: JC4301

C;Accession: JC4301

Gene 163, 257-261, 1995
A:Title: The porcine follitropin receptor: cDNA cloning, functional expression and chromosomal localization
A:Reference number: JC4301; MUID:96011644; PMID:7590277
A:Accession: JC4301
A:Molecule type: mRNA
A:Residues: 1-694 <REM>
A:Cross-references: GB:L31966
A:Experimental source: ovarian granulosa cells
C:Comment: This receptor belongs to the family of the G-protein coupled receptors. It has a role in oogenesis in male and oogenesis in female.
C:Genetics:
A:Gene: fshr
A:Map position: 3 q2.2-q2.3
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein
F:1-365/Domain: follicle-stimulating hormone binding #status predicted <HOB>
F:70-94/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
F:366-388/Domain: transmembrane #status predicted <TM1>
F:398-420/Domain: transmembrane #status predicted <TM2>
F:443-464/Domain: transmembrane #status predicted <TM3>
F:485-507/Domain: transmembrane #status predicted <TM4>
F:528-549/Domain: transmembrane #status predicted <TM5>
F:573-596/Domain: transmembrane #status predicted <TM6>
F:608-629/Domain: transmembrane #status predicted <TM7>

Gene 163, 257-261, 1995


```
QY 673 ETKAPSSKVIILLCALLALTMAAVPLLGSKYKASPLCLPLPGEPESTMGVVALILL 732
Db 477 QCKQVRRHAASIMLYGWIFAFTVALPFIIGISSYMKVSIICLPMDIDSPLSQLYVVSLLVL 536
QY 733 NSLCFLMWTIAVTKLYCNL-DKGDLENTWDCSMVKHIALLLPTNCILNCPVAFLSFSLI 791
Db 537 NVLAFVILGCVIHIYLVTRPNINSSSDYKIAKEMAMLIPTDFLCWVPISFFAISASL 596
QY 792 NLTFISPEVIVKILLVAVVPLPACLMPLLYILPNPHFKEDLVLSKQTYVWTRSKHPSLMS 851
Db 597 KVPLITVSKILLVLFVPIINSCANPFLYFTKFNFRDVFIL----- 639
QY 852 INSDVEKQSCDSTQALVTFSS-SITVDLPSSVPSPAYP-VTESCHLSSVAVFP 905
Db 640 -----LSKFGCYEMOQTYRTENLSTAHNIHPRNGHCPAPRITNS-----SSTLIP 687

RESULT 9
JC7361
follicotropin receptor precursor - new
N:Alternate names: follicle-stimulating hormone receptor
C:Species: Cynops pyrrhogaster (new)
C>Date: 17-Nov-2000 #sequence_revision 17-Nov-2000 #text_change 09-Jul-2004
C:Accession: JC7361
R:Nakayama, Y.; Yamamoto, T.; Oba, Y.; Nagahama, Y.; Abe, S.
Biochem. Biophys. Res. Commun. 275, 121-128, 2000
A:Title: Molecular cloning, functional characterization, and gene expression of a follicle-stimulating hormone receptor
A:Reference number: JC7361
A:Contents: Testis
A:Accession: JC7361
A:Molecule type: mRNA
A:Residues: 1-696 <NA>
A:Cross-references: UNIPROT:Q9DGF5; DBJ:AB005587
C:Comment: This protein, containing seven transmembrane domains and a large glycosylated in and thyroid stimulating hormone. This receptor has a common signal transduction pathway
C:Genetics:
A:Gene: fsh-r
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C:Keywords: disulfide bond; glycolysis; glycoprotein; hormone receptor; testis; transmembrane
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-696/Product: follicle-stimulating hormone receptor #status predicted <MAT>
F:18-359/Domain: extracellular #status predicted <EXT>
F:370-389/Domain: transmembrane #status predicted <TM1>
F:402-424/Domain: transmembrane #status predicted <TM2>
F:447-468/Domain: transmembrane #status predicted <TM3>
F:489-511/Domain: transmembrane #status predicted <TM4>
F:532-553/Domain: transmembrane #status predicted <TM5>
F:577-600/Domain: transmembrane #status predicted <TM6>
F:612-633/Domain: transmembrane #status predicted <TM7>
F:46,190,198,267,292/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:445-520/disulfide bonds: #status predicted

Query Match 13.3%; Score 625; DB 2; Length 696;
Best Local Similarity 22.0%; Pred. No. 1e-32;
Matches 200; Conservative 140; Mismatches 313; Indels 256; Gaps 23;

QY 10 LSLPVLQLATGGSPRGVLLRGCPTRCHCEPDGMRLLVR-DCSDLGLSLPNSLVFT 68
Db 1 MSLAILCLLLAVGSS-----FGCHPVCR-----LNRVFTQESHVQIPR----- 41
QY 69 SYLDSMNINISQLLPNPLPSLRFLEELRAGNALYIPKGAFTGLYSIK-VLMQLNNQLR 127
Db 42 ---DIPRNS-----TELRFVLTKVTIPKAAFGSGFEDVENIEISQNDVLK 83
QY 128 HVPTEALQNLSLQSLRLD-ANHISYVPSFCFSGHLSRLHLLDNLALTEIPVQAFRLS 186
Db 84 TIEANVFSHLPKLRIRIEKANNLVYIDPFAQNLPSLKLILNLTGTLQVLP----- 135
QY 187 ALQAMTLAKNIHHPDVAFGNLSLVVHLHNN-RIHSLGKCKCDGSHSLLETLDLYNN 245
Db 136 -----AVSKIRSFHS-----VLLDVQDNINIRHIGKNSFAGLSS-ESTIRLN- 177
QY 246 LDEFPPTAIRTLNLKELGFHGNINRSIPEKAFVGNPNSLITTHFYDNPNIQFVGRSAFQHL 305
```

```
Db 178 -----KNGIEEQNHAFNGT-----HUN 195
QY 306 ELRTLTLNGASQITEFFDLTGTANLESULTGAQISSLPQTVCNQLPNQLVLDLSTYNLLE 365
Db 196 E-----LNLSDN----- 202
QY 366 DLPFSVCQKLOKIDLRNEIYKVDTPFOQLLSRLSNLAWNKIAIHPNNAFSLPSLI 425
Db 203 -----ORLEKLP-----DQVFOGATGVP 220
QY 426 KLDLSSNLLSFPITGLHGLTHLKLGTGNHALQSLISSNEPELKVIEPYPAYQCCAFG-- 483
Db 221 ILDISRTRIHFLPNNGIENIKKFRARFNYILKKLPPELEKFAELIEANLITYPSHCAPANR 280
QY 484 -----VCENAYKISQNWNGKDNSSMDLHKDKAGMFOAQDRDLDFLLDFFED-- 532
Db 281 ERKSEMHPICNKSF-----GKHDSAEKPEDKQLRRFSNEDYLSYSGFSYLVENG 332
QY 533 -----LKALHSVQCSPPSPGPKPCHEHLLDGLMIRIGVMTIAVLATCNALVTSTVPR 584
Db 333 EFNVDYILCNEVHDVICFPKPDAPNCPEDIMGDNTRVLIMLISILAITGNITVLVILIS 392
QY 585 SPLYSIPKLIGVIAAVNMLTVSSAVLAGVDAFTFGSFARHGAWENGVCCHVIGPLS 644
Db 393 SOYKFTVPRFLMCLNAPADLCMGYLLLASVDIKTSQYNNHAIQWOTGSCAAAGFT 452
QY 645 IFASESSVFLITLAALBERGFSVKYSAKPEKAPSSSLKVIILLCALLALTMAAVPLLAGS 704
Db 453 VFASLSVYTLTVITIERWHITTYAMQLDKRVFRHATAIWWGVIFAFTVAILPFGVS 512
QY 705 KYGASPLCLPFPGEPESTMGVVALIILNSLCFLMWTIAVTKLYCNL-DKGDLENIWDCS 763
Db 513 SYTKVSIKLPMDIESRLSSQSYIVFLVNLVCAFLIICACYIGYLVTRPNVSVSSDTK 572
QY 764 MVKHIALLLFTNCILNCPVAFLSFSSILNLFISPEVIVKILLVAVVPLPACLMPLLYILF 823
Db 573 IAKEMALIFTDFLCMAPISEFAISAKLIKPLITVSKSKILLVLPYINSCANPFLYAF 632
QY 824 NPHEKEDLVLSLRKQTYVWTRSKHPSLMSINDDVEKQSCDSTQALVTFSSSI--TYDLP 881
Db 633 TKTFRRDFIL-----MSKFGCEVQAQNCR-----TETSSSLHSHWRN 672
QY 882 PSSVSPSPAY 890
Db 673 GHYVPAPKY 681
```

RESULT 10

QKHUFT

follicotropin receptor precursor - human

N:Alternate names: follicle stimulating hormone receptor (FSHR)

N:Contents: follicle-stimulating hormone precursor long splice form; follicotropin receptor precursor

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1991 #sequence_revision 06-Sep-1996 #text_change 09-Jul-2004

C:Accession: 157661; I56448; PC1147; S30560; I57672; JN0122

R:Gromoll, J.; Dankbar, B.; Gudermann, T.

Mol. Cell. Endocrinol. 102, 93-102, 1994

A:Title: Characterization of the 5' flanking region of the human follicle-stimulating ho

A:Reference number: 157661; MUID:95011044; PMID:7926278

A:Accession: 157661

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-51 <GRO>

A:Cross-references: UNIPROT:P23945; GB:S73199; NID:g685036; PIDN:AAB32071.1; PID:g685037

R:Gromoll, J.; Ried, T.; Holtgreve-Grez, H.; Nieschlag, E.; Gudermann, T.

J. Mol. Endocrinol. 12, 265-271, 1994

A:Title: Localization of the human FSH receptor to chromosome 2 p21 using a genomic prob

A:Reference number: 156448; MUID:95000244; PMID:7916967

A:Accession: 156448

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 286-695 <GR2>

A;Cross-references: GB:S73526; NID:g688069; PIDN:AA32225.1; PID:g688070
R;Gromoll, J.; Gudermann, T.; Nieschlag, E.
Biochem. Biophys. Res. Commun. 186, 1077-1083, 1992
A;Title: Molecular cloning of a truncated isoform of the human follicle stimulating hormone
A;Reference number: PC1147; MUID:93075197; PMID:1359889
A;Accession: PC1147
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-223,286-294, 'P', 296-342 <GR3>
A;Cross-references: EMBL:X68044; NID:g31473; PIDN:CAA48179.1; PID:g31474
A;Experimental source: testis
R;Gromoll, J.
submitted to the EMBL Data Library, August 1992
A;Reference number: S30560
A;Accession: S30560
A;Molecule type: mRNA
A;Residues: 1-12, 'R', 14-223,286-294, 'P', 296-342 <GR4>
A;Cross-references: EMBL:X68044; NID:g31473; PIDN:CAA48179.1; PID:g31474
R;Kelton, C.A.; Cheng, S.V.; Nugent, N.P.; Schweickhardt, R.L.; Rosenthal, J.L.; Overton
Mol. Cell. Endocrinol. 89, 141-151, 1992
A;Title: The cloning of the human follicle stimulating hormone receptor and its expression
A;Reference number: 157672; MUID:93246012; PMID:1301382
A;Accession: 157672
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-679, 'N', 681-695 <KEL>
A;Cross-references: GB:S59900; NID:g300072; PIDN:AA26480.1; PID:g300073
R;Minagishi, T.; Nakamura, K.; Takakura, Y.; Ibuki, Y.; Igarashi, M.
Biochem. Biophys. Res. Commun. 175, 1125-1130, 1991
A;Title: Cloning and sequencing of human FSH receptor cDNA.
A;Reference number: JN0122; MUID:91222171; PMID:1709010
A;Accession: JN0122
A;Molecule type: mRNA
A;Residues: 1-111, 'T', 113-136, 'AV', 199-306, 'A', 308-695 <MTN>
A;Cross-references: EMBL:M65085; NID:g182770; PIDN:AA52477.1; PID:g182771
C;Genetics:
A;Gene: GDB:FSHR
A;Cross-references: GDB:127510; OMIM:136435
A;Map position: 2p21-2p16
A;Introns: 223/3
A;Note: the exact position of the intron cannot be determined from the experimental data
C;Function:
A;Description: receptor that mediates the biochemical effects of follitropin
A;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
F;1-695/Product: follitropin receptor precursor, long splice form #status predicted <SPH
F;1-223,286-695/Product: follitropin receptor precursor, short splice form #status predi
F;1-15/Domain: signal sequence #status predicted <SIG>
F;16-695/Product: follitropin receptor #status predicted <MAT>
F;16-366/Domain: extracellular hormone binding #status predicted <EHB>
F;56-70/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F;71-95/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F;96-120/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F;121-145/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F;146-169/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
F;172-193/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
F;194-218/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
F;367-387/Domain: transmembrane #status predicted <TM1>
F;398-421/Domain: transmembrane #status predicted <TM2>
F;444-465/Domain: transmembrane #status predicted <TM3>
F;486-508/Domain: transmembrane #status predicted <TM4>
F;529-550/Domain: transmembrane #status predicted <TM5>
F;574-597/Domain: transmembrane #status predicted <TM6>
F;609-630/Domain: transmembrane #status predicted <TM7>
F;191,199,293,318/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;555/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
F;596/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted

Query Match 13.2%; Score 621.5; DB 1; Length 695;
Best Local Similarity 22.4%; Pred. No. 1.7e-32;
Matches 207; Conservative 146; Mismatches 36; Indels 265; Gaps 25;
8 VLLSLPVLQLATGSGSPRSVLLRGCPHCH---CEPDGRMLLRVDCSLGLSELPNL 64

Db 4 LLVSLLAFLSLGSG-----CHRHCHCSNRVFL---CQBSKVTEIPS-- 42
Qy 65 SVFTSYDLDLNNNISQLLPNPLSLRFLERLRLAGNALTYIPKGAFTGLVSL-KVLMLO 123
Db 43 -----DLPRNAI-----ELRFVLTKLRLVIQKGAFGSGFGLERIEISQN 80
Qy 124 NQLRHVTEALQNLRLSLOSLRLD-ANHISVVPSPCSGLHSLRLHLWLDNLTALTEIPVQAF 182
Db 81 DVLVEIEADVFSNLPKLUHEIRIEKANNLLYNP-----EAF 116
Qy 183 RSLSLAQAMTLALNKKIHHIPDYAFAGNLSLVVLHLHNNRIHSLGKCGFDGLHSLLETLDLN 242
Db 117 QNLNQLQYLLISNTGIKHLPDV-----HKHLSLQK-----VLLDIQ 152
Qy 243 YNNLDEPPTAIRLSNLKELGFSNNNIRSIPEKAFVGNPSLIITHFDVNPQIOFVGRSAFO 302
Db 153 DN-----INHTIERNISFVGLS-----FESVILWLNKNGIQ 183
Qy 303 HLPRLRLTLNGASQITEFPDLTCTANLESITLTGAQISSLPQVNCNLPNLQVLDLSYN 362
Db 184 ---EIHNCAPNG-----TOLDELNUSD--NN 204
Qy 363 LLEDLPFSVCQKLIKIDLRHNEIYEIKVDTFQOLLSLRSLNLAWNKIAIHPNASTLP 422
Db 205 NLEELPN-----DVFHGAS 218
Qy 423 SLIKLDLSSNLLSPFTTGLHGLTHLKLGTNHALQSLISSENPELKVEMPYAYOCCAF 482
Db 219 GPVILDISRTIRIHSPLSYGLNKLRAARSTYNLKKLPLEKLVALMEASITSPSHCCAF 278
Qy 483 G-----VCENAYKISNQ-----WNKGDNSMDDLHKDKAGFQAQDRDLDF 525
Db 279 ANWRQISELHPICNKS--ILROEVDVMTQGRQSSLAEDN-----ESSYRGDFMT 329
Qy 526 LLDPEEDL-KALHVSQCSPPGPKCEHLLDGLWLIIRIGVMTIAVLATCNALVTSTVFR 584
Db 330 YTEFDYDLCEVVDVTCSPKPDAPNCPEDIMGYNILAVLIWFISILAITGNIIVLVILT 389
Qy 585 SPLYISPIKLLIGVIAVNMMLTGVSVAVLGADVFTFGSFARHGAWNGVGVCHVIGLS 644
Db 390 SQYKLVPRFLMCLNAPLADLCIGIYLLLIASVDIHTKSQYHNYAIDMQTAGCDAAGPFT 449
Qy 645 IFASESSVFLTLAALRGESVKYSKAFEPKAPSSLLKVIILCALLALTAAPVPLGGS 704
Db 450 VFASLSVYVLTALTLERWHTITHAMQDCKVQLRHASVVMWGNIFAFAPALPFIIGIS 509
Qy 705 KYGASPLCLPLPFGPESTMGMVALILLNSLCFLMMTIAVTKLYCNLDKGLD-ENIWDCS 763
Db 510 SYMKVSICLPMDIDSPLSQLYVMSLVNLVLAFFVICGCIHIYLTVRNPNIVSSSDTR 569
Qy 764 MVKHIALLLFTNCILNCPVAPLSPSSLINITFTISPEVIKILLVVVLPDACLNPILLILF 823
Db 570 IAKRMAMLIPTDFLCMAPISFPAISLKVPLITVSKAKILLVLFHPINSCANPLFAIF 629
Qy 824 NPHEKEDLVSLRKQTYVWTKSKHPSLMSINSDDVKEQSCDSTQALV--TFTSSSITVDLP 881
Db 630 TKNFRRRFFIL-----LSKCGCYEMQAQIYRTTSTSTVHNTHP 667
Qy 882 PSSVPSPAYPVTERSCHLSSVAVFP 905
Db 668 RNHCSSAPRVTSQ---STYLVP 688

RESULT 11

A42395
lutropin receptor - mouse
N;Alternate names: luteinizing hormone-choriogonadotropin receptor
C;Species: Mus musculus (house mouse)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: A42395
R;Gudermann, T.; Birnbaumer, M.; Birnbaumer, L.
J. Biol. Chem. 267, 4479-4488, 1992

A>Title: Evidence for dual coupling of the murine luteinizing hormone receptor to adenylyl
mone receptor expressed in L cells.

A;Reference number: A42395, MUID:92165799; PMID:1311310

A;Accession: A42395

A;Status: Preliminary

A;Molecule type: mRNA

A;Residues: 1-700 <GUD>

A;Cross-references: UNIPROT:P30730; GB:M81310; GB:M81318; NID:g198811; PIDN:AAA39432.1;

A;Note: sequence extracted from NCBI backbone (NCBIN:84064, NCBIP:84066)

C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

C;Keywords: G protein-coupled receptor; transmembrane protein

F;54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F;78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F;103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F;128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

F;180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>

F;202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>

Query Match 12.8%; Score 600; DB 2; Length 700;
Best Local Similarity 23.0%; Pred. No. 4.3e-31;
Matches 207; Conservative 143; Mismatches 320; Gaps 23;

QY 5 RLGVLISLPVLIQATGSSPRSGVLLRG-CPTHCHCEPDGEMLRVDCSLGSELPSN 63

DB 4 RVPALRQLLVLAUMLVKOSQLHSPSLGSRCEPCDCAPDG--ALRCPGPRAGLRL--- 58

QY 64 LSVFTSYLDLSNMNITQLLENPLSLRFLLELRAGNALTYIPKGAFTGLYS-LKVLMLQ 122

DB 59 -----SLTYL-----PVKIPSPQAFGLNEVVKIESQ 86

QY 123 NNQLRHVTEALQNLRSQSLRL-DANHI SYVPPSCFGLSLRHLMLDNDNALTEIPVQA 181

DB 87 SDSLERIEANAFDNLNLSEILIQNTKLLYIEPCGFTNLPRLYLSI----- 134

QY 182 FRSLSAQMTLAKMIIHPIDYAFGNLSLVLLHNNRHSLSLKKKCPDGLHSLETLDL 241

DB 135 -----CMTGIRTLF--DV 145

QY 242 NVNLDDEPTAIRTLSNLKELGFHNNITRISPEKAFVG-NPSLTIHFVDNPIQVGRSA 300

DB 146 SKISSSENFLEICDNL-----YITIPGNAFQGNNEISITKLYGNGFEVQSHA 197

QY 301 FOHLPELRTLTLNGASQITEFPDLTGTLANLESLLTGAQISLLPQTCNQLPNLQVLDLS 360

DB 198 F-----NGTTLS-----LELK 209

QY 361 YNL-LEDLPSPSVCKLOKIDLRHNEIYEIKVDTFQQLLSRLSLNLANWKIATIHNPAPS 419

DB 210 ENIVLEKMH-----GTFQG-----A 225

QY 420 TLPSLTKLDLSSNLLSSPITGLHGLTHLKLGNHALQSLISSENFPPELKVTEMPYAYQC 479

DB 226 TGPSI--LDVSTKQLQALPSHGLESIQTLATSYSLKTLPSREKFTSLLVATLYTPSHC 283

QY 480 CAFGVCE-----AYKISQNMKNQNSMDLHKD--AGMFQADRDLEDLDFEED 532

DB 284 CAFRNLPKKEQNFSSIFENFSKQCESTVREANNETLYSAIF--EENELSGW--DYDYD 338

QY 533 LKALHSVQCSPPGPKCEHLLDGLWLRIGVWTTAVLATCNALVTSTVFRSPLYISPI 592

DB 339 FCSPTKLOCTPEPDAPNCPEDIMGYAFRLVILWILNLAIFGNLVFLVLLTSRYKLVTP 398

QY 593 KLLIGVIAAVMLTCVSSAVLAGVDAFTFGSPARHGWNGVGVCHVIGLSIFASESSV 652

DB 399 RFLMNLCSFADFCMGLYLLLTASVDSTQKQYNNHAIQWGTGSGCSAAGFTVFASELSV 458

QY 653 FILLTALAALRGFSVKYSKAFETKAPFSSLKVILLICALLATMAAVPLLGSKYCASPLC 712

DB 459 YLTIVTLERWHITTYAVQLDQKLRLRAIPMLGWIIFSTLMATLPLGVSSYMKVVIC 518

QY 713 LPLPFGEFTWGMVALILLNSLCFLMTMTIATYKLYCNLDKGL-ENTWDCSMVKHALL 771

Db 519 LPMDEVSTLSQVYILSILLNNAFVVICACVRIYFAVQNPDELTPAKNDTKIAKQMAIL 578

QY 772 LFTNCILNCPVAFISFSSLINLFTISPEVIRFILLVVPVPLPACLNPLLYILFNPHFKBDL 831

Db 579 IFTDPTCMAPISFPAISAAFKVPLITVTNSKVLVLFYVNSCANPELYAVFTKAFQDF 638

QY 832 VSL-----RKQTYVWTSKSPSLMSINSDDVEKSCDSTQALVTFTSSSIYDLP 882

Db 639 PLLLSRFGCCKHRAELY---RRKEFSACTFNSKNGFPRSSKPSQAALKLSIVHCQOQTPP 695

RESULT 12

QYHURH

thyrotropin receptor precursor - human

N;Alternate names: thyroid stimulatory hormone receptor (TSHR)

C;Species: Homo sapiens (man)

C;Date: 22-Jan-1993 #sequence revision 01-Mar-1996 #text_change 09-Jul-2004

C;Accession: A33789; A33786; A34052; A36120; S38280

R;Libert, F.; Lefort, A.; Gerard, C.; Parmentier, M.; Perret, J.; Ludgate, M.; Dumont, J.

Biochem. Biophys. Res. Commun. 165, 1250-1255, 1989

A;Title: Cloning, sequencing and expression of the human thyrotropin (TSH) receptor: evi

A;Reference number: A33789; MUID:90121232; PMID:2610690

A;Accession: A33789

A;Status: nucleic acid sequence not shown; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-86, 'L', '88-764 <LIB>

A;Cross-references: UNIPROT:P16473

R;Nagayama, Y.; Kaufman, K.D.; Seto, P.; Rapoport, B.

Biochem. Biophys. Res. Commun. 165, 1184-1190, 1989

A;Title: Molecular cloning, sequence and functional expression of the cDNA for the human

A;Reference number: A33786; MUID:90121223; PMID:2558651

A;Accession: A33786

A;Molecule type: mRNA

A;Residues: 1-600, 'H', '602-764 <NAG>

A;Cross-references: GB:M31774; NID:G340003; PIDN:AAA36783.1; PID:G340004

R;Misrahi, M.; Loofofelt, H.; Alger, M.; Sar, S.; Guloichon-Mantel, A.; Milgrom, E.

Biochem. Biophys. Res. Commun. 166, 394-403, 1990

A;Title: Cloning, sequencing and expression of human TSH receptor.

A;Reference number: A34052; MUID:90147730; PMID:2302212

A;Accession: A34052

A;Molecule type: mRNA

A;Residues: 1-726, 'E', '728-743, 'K', '745-764 <MIS>

A;Cross-references: GB:M32215; NID:G307524; PIDN:AAA61236.1; PID:G307525

A;Note: 13-Pro, 260-Pro, 414-His, 500-Leu, 634-Leu, 727-Asp, and 744-Asn were also found

R;Frazier, A.L.; Robbins, L.S.; Stork, P.J.; Sprengel, R.; Segaloff, D.L.; Cone, R.D.

Mol. Endocrinol. 4, 1264-1276, 1990

A;Title: Isolation of TSH and LH/CG receptor cDNAs from human thyroid: regulation by tis

A;Reference number: A36120; MUID:91155962; PMID:2293030

A;Accession: A36120

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-129, 'S', '131, 'AFS', '135-195, 'D', '197, 'F', '199-256, 'S', '258-263, 'A', '265-305, 'IET

A;Cross-references: GB:M73747; NID:G903759; PIDN:AAA70232.1; PID:G903760

R;Parma, J.; Duprez, L.; van Sande, J.; Cochaux, P.; Gervy, C.; Mockel, J.; Dumont, J.;

Nature 365, 649-651, 1993

A;Title: Somatic mutations in the thyrotropin receptor gene cause hyperfunctioning thyroc

A;Reference number: S38280; MUID:94019814; PMID:8413627

A;Accession: S38280

A;Molecule type: DNA

A;Residues: 615-642 <PAR>

A;Experimental source: thyroid adenomas

A;Note: mutations to Cys-619 or to Ile-623 in hyperfunctioning thyroid adenomas cause the

C;Comment: See PIR:JC1319 for a splice form of thyrotropin receptor which lacks the sevc

C;Genetics:

A;Gene: GDB:TSHR

A;Cross-references: GDB:125313; OMIM:275200

A;Map position: 14q31-14q31

A;Introns: 231/2

A;Note: the list of introns is incomplete; the eighth of eight introns is shown

C;Function:

A;Description: receptor that mediates the biochemical effects of thyrotropin

C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h

C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece

Db 88 -----TLORLSHSEFYNSKWMHIEIRNTR-----SLTSD----- 118
Qy 248 EFPPTAIRTLNKLKELGPHNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSAFOHLPEL 307
Db 119 --PDALKELPLKPLGIGTNGVFPDVKV-----YSTDVFI----- 155
Qy 308 RLTTLNGASQITEFPDLTGANLESLLTGAQISLP-----QTVCNQLPNQVLDLSYNL 363
Db 156 -----LEITDNP-----YMASIPANAFQGLCNETLTKL-----YN- 186
Qy 364 LEDLPFSVSCQKQKIDLRHNEIYKVDYTFQOLLSRLSLANWK-TAIHPNAPFTL- 421
Db 187 -----NGFTSQGHAFNG-TKLDVYLNKKNKYSATDKDAFGVY 225
Qy 422 --PSLIKLDLSNLSPPITGLHGLTHLKTGNHALQSLISSENPPELKIEMPIYQOC 479
Db 226 SGPTL--LDVSVTSYVTLPSKGLHKLIELARNTWTAKKPLSLSLHLTRADLSYPSHC 283
Qy 480 CAP-----GVCE-----NAYKISNOWNKGDNSMDLHK--KDAQMP 514
Db 284 CAFKNQKIRGILESMLCNESIRSLQRKSVNTLNGPFDQOEYELGDSHAGYKDSQF 343
Qy 515 QAQDE-----RLEDFLDF-----BEDLKALHS-----VOCSPSPG 546
Db 344 QOTDSNHYVFEQEDEILGFGQELKNPOBELQAFDSDHYDVTVCNGNEDMVCPTKSD 403
Qy 547 PFKPCEHLLDGWLIRIGVTTAVALTCTNALVTSTVFRSPLYISPIKLLIGVIAVNMLT 606
Db 404 EFNPCEDIMGYKFLRIVVWFVSLALLGNVFLVLLTSHYKLTVPFLMCLNAPADFCM 463
Qy 607 GVSSAVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFASESSVFLTLAALRGFSV 666
Db 464 GMYLLIASVDLYTHSEYNNHAIQDTGPGCNTAGFTVFASLSVYTLTITLERWYAI 523
Qy 667 KYSAFETKAPFSSLVII-----LLCALLATMAAVPILGSKYKASPLCLPLPGEPEST 722
Db 524 TFAMRLDRIRURHAYAVMGWGWCCFLAL-----LPLVGISSYKAVKVICLPMDETETPLA 579
Qy 723 MGYVVALILNLSCLPMMTIATYTKLYCNL-----DKGDLNIWDCSMVKHIALLLFTNCI 777
Db 580 LAYIILVLLNIVAIIVCSCVKIYITVRNPQYNGDK-----DTKIARMAVLIFTDFM 635
Qy 778 LNCVPVAFSFLSINLITISPEVHKFILLVVVPLPACLNPLLYILFNPHFKEDLVSL--- 834
Db 636 CMAPISFYALSALNMKPLITVNSKILLVLFVPLNSCANPFLYAFITKAFORDVPEILSK 695
Qy 835 ----RKQTYVWTRSKHPSLMSINSDDEKQSCDSTQALVTPTSSSITDYDLPSSVSPAY 890
Db 696 FGICKRQAOAY-RGORVSPKNSAGIQIKVTRDMRQSLPNQDE-----Y 739
Qy 891 PVTESCHLS 899
Db 740 ELLENSHLT 748
RESULT 14
177463
luteinizing hormone/chorionic gonadotropin receptor - rat
C:Species: Rattus sp. (rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C:Accession: 177463
R:Aatsinki, J.T.; Pietila, E.M.; Lakkakorpi, J.T.; Rajaniemi, H.J.
Mol. Cell. Endocrinol. 84, 127-135, 1992
A:Title: Expression of the LH/CG receptor gene in rat ovarian tissue is regulated by an
A:Reference number: 15768; PMID:92347604; PMID:1353463
A:Accession: 177463
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-700 <RES>
A:Cross-references: GB:S40803; NID:g252163; PIDN:AAB22680.1; PID:g252164
C:Genetics:
A:Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
C:Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat H

F:153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>
Query Match 12.5%; Score 588.5; DB 2; Length 700;
Best Local Similarity 22.6%; Pred. No. 2.4e-30;
Matches 205; Conservative 143; Mismatches 309; Indels 251; Gaps 25;
Qy 6 LGVLLSPVLQOLATGSSPSRSGVLLRCPCPHCHCEPDGRMLLRVDCSDGLGSELPSNLS 65
Db 14 LAVILLKPSQLO-----SRELSG---SRCPEPCADPG--ALRCPPGRAGLARL----- 58
Qy 66 VFTSVLDLSMNNISQLLPNPLSRFLBELRLAGNALTYIPKGAFTGLYS-LKVLMLQNN 124
Db 59 -----SITYL-----PVKVIPOAERGLNEVVKIISQSD 88
Qy 125 QLRHVPTEALQNLRSLOSRL-DANHSYVPPSCFSGHLSRHLWLDNALTEIPVQAFR 183
Db 89 SLERIANAPDNLANLSELLONTKNLLYIEFGAFTNLPRLKYLSI----- 134
Qy 184 SLSALQMTLALNKIHHIPDYAFGNLSLVVLHNNRIHSLGKKCFDGLHSLTDLNY 243
Db 135 -----CNTGIETLP--DVTK 147
Qy 244 NNLDEPTAIRTLNKLKELGPHNNIRSIPEKAFVG-NPSLTIHFYDNP IQVGRSAFO 302
Db 148 ISSSEFNILICDNL-----HITIPGNAFOGMNESVTCLKYNGFEVQSHAF- 198
Qy 303 HLPBLRTLTNGASQITEFPDLTGANLESITLTGAQISLPQTVCNQLPNQVLDLSYN 362
Db 199 -----NGTTLIS----- 205
Qy 363 LLEDLPFSVSCQKQKIDLRHNEIYKVDYTFQOLLSRLSLANWKTAIHPNAP--ST 420
Db 206 -----LELKEN-IYLEK-----MHSGAFOGAT 226
Qy 421 LPSLIKLDLSNLSPPITGLHGLTHLKTGNHALQSLISSENPPELKIEMPIYAOCC 480
Db 227 GPSI--LDISSTKQALPSHGLSIOQLIALSSVSLKTLPSKEKFTSLVATLTPSHCC 284
Qy 481 A-----FGVCENAYK-ISBNOWNKGDNSMDLHKDKAGMFOAQDERDLEDFIL 527
Db 285 AFRNLPKKEQNFSPSIPENFSKQCESTVRKADNETL-----YSAIF--EENELSGW-- 333
Qy 528 DFEEDLKALHSVQCSPPSPGPKPCHEHLLDGLWLRIGVWTTIAVLATCTNALVTSTVFRSPL 587
Db 334 DYDYGFCSPKTLQCAPEDAFNPCEDINGYAFRLVLIWLINILAI FGNLT VLVLLTSRY 393
Qy 588 YISPIKLLIGVIAAVNMLTGVSSAVLAGVDAFTGSPARHGAWENGVGCHVIGFLSIFA 647
Db 394 KLTVPRLMCLNSFADFCMGYLLLIASVDSQTKQYTNHAIQDTGSGCGAAGFTVFA 453
Qy 648 SESSVFLITLAALBERGFSVKYSAPETKAPFSSLSKVIILLCALLATMAAVPLLGSKY 707
Db 454 SELSVYTLTVITLERWHITTYAVQDDQKLRLHAIPMLGGWLFSTLIATPLVGSINYM 513
Qy 708 ASPLCLPLPGEPESTMGVNWALILNLSCLPLMTIATYTKLYCNLDKGLD-ENIWDSCSVK 766
Db 514 KVSICLPMDEVTLSQVYILSILINLVAVFVVCACYIRIYFAVQNPBELTAPNKDTKAK 573
Qy 767 HIALLTNCLNCPVAFSPSLNLTTFISPEVHKFILLVVVPLPACLNPLLYILFNPH 826
Db 574 KMAILIFTFTCMAPISFAISAFKVPPLITVTSKILLVLFVYVNSCANPFLYAFITKA 633
Qy 827 FKEDLVSL-----RKQTYVWTRSKHPSLMSINSDDEKQSCDSTQALVTPTSSSIT 877
Db 634 FORDFLLLSRFGCKRAELY---RRKEFSAYTSNCKNGPFGASKPKSQATLKLSTVHCQ 690
Qy 878 YDLPPSSV 885
Db 691 QPIPPRAL 698
RESULT 15
A49744

lutropin-choriogonadotropin receptor precursor - rat
 N;Alternate names: luteinizing hormone-choriogonadotropin receptor
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 03-May-1994 #sequence_revision 13-Sep-1998 #text_change 09-Jul-2004
 C;Accession: A49744; A40545; A41343; A61453; A32460
 R;Koo, Y.B.; Ji, I.; Slaughter, R.G.; Ji, T.H.
 Endocrinology 128, 2297-2308, 1991
 A;Title: Structure of the luteinizing hormone receptor gene and multiple exons of the c
 A;Reference number: A49744; MUID:91209270; PMID:2019252
 A;Accession: A49744
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-700 <KOO>
 A;Cross-references: UNIPROT:P16235; GB:M68917
 A;Note: authors translated the codon CAA for residue 307 as Glu; AAC for residue 355 as
 R;Tsai-Morris, C.H.; Buczko, E.; Wang, W.; Xie, X.Z.; Dufau, M.L.
 J. Biol. Chem. 266, 11355-11359, 1991
 A;Title: Structural organization of the rat luteinizing hormone (LH) receptor gene.
 Science 245, 494-499, 1989
 A;Reference number: A40545; MUID:91250455; PMID:2040640
 A;Accession: A40545
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-320 <TSA>
 A;Cross-references: GB:M63918; GB:M63919; GB:M63920; GB:M63921; GB:M63922; GB:M63923; GB
 R;McFarland, K.C.; Sprengel, R.; Phillips, H.S.; Koehler, M.; Roseblit, N.; Nikolics, K
 Science 245, 494-499, 1989
 A;Title: Lutropin-choriogonadotropin receptor: an unusual member of the G protein-couple
 A;Reference number: A41343; MUID:89332512; PMID:2502842
 A;Accession: A41343
 A;Molecule type: mRNA
 A;Residues: 1-700 <MCF>
 A;Cross-references: GB:M26199; NID:g205178; PIDN:AAA41528.1; PID:g205179
 R;Dufau, M.L.; Minegishi, T.; Buczko, E.S.; Delgado, C.J.; Zhang, R.
 J. Steroid Biochem. 33, 715-720, 1989
 A;Title: Characterization and structure of ovarian and testicular LH/hCG receptors.
 A;Reference number: A61453; MUID:90097014; PMID:2601325
 A;Accession: A61453
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 27-33, 'X', 35-37, 'X', 39, 'X', 41-44 <DUF>
 R;Roche, P.C.; Ryan, R.J.
 J. Biol. Chem. 264, 4636-4641, 1989
 A;Title: Purification, characterization, and amino-terminal sequence of rat ovarian rece
 A;Reference number: A32460; MUID:89174723; PMID:2925659
 A;Accession: A32460
 A;Molecule type: protein
 A;Residues: 27-32, 'LX', 35-37 <ROC>
 C;Genetics:
 A;Introns: 58/2; 82/2; 107/2; 132/2; 157/2; 183/2; 206/2; 231/2; 293/2; 320/2
 C;Superfamily: glycoprotein hormone receptor; leucine-rich alpha-2-glycoprotein repeat h
 C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; hormone rece
 F;54-77/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
 F;78-102/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
 F;103-127/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
 F;128-152/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
 F;153-177/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>
 F;180-201/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR6>
 F;202-226/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR7>
 F;368-389/Domain: transmembrane #status predicted <TM1>
 F;400-422/Domain: transmembrane #status predicted <TM2>
 F;444-466/Domain: transmembrane #status predicted <TM3>
 F;489-511/Domain: transmembrane #status predicted <TM4>
 F;530-551/Domain: transmembrane #status predicted <TM5>
 F;575-598/Domain: transmembrane #status predicted <TM6>
 F;610-631/Domain: transmembrane #status predicted <TM7>
 F;103,178,199,295,303,317/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;647,648/Binding site: palmitate (Cys) (covalent) #status predicted
 F;681/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted

Query Match 12.5%; Score 588.5; DB 2; Length 700;

Best Local Similarity 22.6%; Pred. No. 2.4e-30;

Matches 205; Conservative 143; Mismatches 309; Indels 251; Gaps 25;

Search completed: July 12, 2005, 07:38:33
 Job time : 47 secs

Qy 6 LGVLLSLPVLQLATGSSPRSGVLLRGCPHCHCEPDGRMLLRVDCSDLGSLSPNLS 65
 |||||
 Db 14 LAVULLPSQIQ-----SRELGG---SRCEPCDCAPDG--ALACPFGRAGLARL----- 58
 Qy 66 VFTSYLDLSMNNSQLLPNPLPSLRFLEBLRAGNALTYIPKGAFTHGLYS-LKVLMLQNN 124
 |||||
 Db 59 -----SLTYL-----PVKVIPOQAQFRLGNLVVVKIEISQSD 88
 Qy 125 QLRHVPTEALQNLSLQSLRL-DANHISYVPPSCFSGLSHSLRHLLWLDNALTEIPVQAFR 183
 |||||
 Db 89 SLERIEANAFDNLNLSELLIQTNTLLYTPGAFTNLPRLKYLSTI----- 134
 Qy 184 SLSALQAMTLALNKIHHPDYAFGNLSSLVVLHNNRIHSLGKKKCDGLHSLFETLDLNY 243
 |||||
 Db 135 -----CNTGRTLP--DVTK 147
 Qy 244 NNLDPEPTAIRTSLNKLGFHNNIRKSIPEKAPVG-NPSLIITHFYDNPQIVGVSAPQ 302
 |||||
 Db 148 ISSSEFNFILEICDNL-----HITTPGNAPQGMNNSVTLKLYGNGFEEVQSHAF- 198
 Qy 303 HLPRLTLTLNGASQITEFPDLTGTANLESILTGAQISSLPQIVCNQLPNQLVLDLSYN 362
 |||||
 Db 199 -----NGTTLIS----- 205
 Qy 363 LLEDLPFSVCQKLOKIDLRHNEIYEIKVDTFQOLLSLRSINLAWNKIAIHPNAF--ST 420
 |||||
 Db 206 -----LELKEN-IYLEK-----MHSGAFQOGAT 226
 Qy 421 LPSLIKLDLSNLSPPITGLHGLTHLKTGNHALQSLISSENFPELKVIEPMYAOCC 480
 |||||
 Db 227 GPSTI--LDISSTKLQALPSHGLSEIQTLLALSSYSKLTLPSEKFTSLVATLTPSHCC 284
 Qy 481 A-----PGVCENAYK-ISNOWNGDDSSMDLHKDAGMFOQADDERLDEPLL 527
 |||||
 Db 285 AFRNLPKKEQNFSPISFENFSKQCESTVRKADNETL-----YSAIF---EENELSGW-- 333
 Qy 528 DFEDLKALHSVQCSPPSPGPKCEHLLDGLIRIGVWTIAVLATCNALVTSTVFRSPL 587
 |||||
 Db 334 DYDYGFCSPKTLQCAPEPDAFNFCEDIMGYAFRLVLIWLIINILAI FGNLTVL FVLLTSRY 393
 Qy 588 YISPIKLLIGVIAAVNMLTGVSSAVLAGVDATFTGSPFARGHAWNGVCHVIGFLSIPA 647
 |||||
 Db 394 KLTVPRLMCLNSPADCMGLYLILLIASVDSQTKGYNNHAIDWQGTSGCGAGFFTVFA 453
 Qy 648 SESVFLTLAALBERGFSVKYSAKFETKAPFSSKLVILLCALIALTMAAVPVLGGSKYG 707
 |||||
 Db 454 SELSVYTLTVITLWHTTITTYAVQLDOKLRHRAIPMLGWLFEFTLIATMPLVIGISNYM 513
 Qy 708 ASPLCLPLPFGEFSTMGYVALILLNSLCFLMMTIATYTKLYCNLDKGL-ENIWDCSMVK 766
 |||||
 Db 514 KVSICLPMDEVSTLSQVYILSILINNVAVFVICACVIRIVFAVQNPETLAPNKDKTKIAK 573
 Qy 767 HIALLLPTNCILNCVAFLSFSSLINLTFISPEVIKILLVVVPLPACLAPLLYILFNPH 826
 |||||
 Db 574 KMAILIFTDTCMAPISEFFAISAAFKVPLITVTNLSKILLVLFYVWNSCANPFLVAIFTKA 633
 Qy 827 FKEDLVSL-----RKQTYVWTRSKHPSLMSINSDVKEQSCDSTQALVTFSSIT 877
 |||||
 Db 634 FQDFLLLLSRFGCKKRAELY---RRKEFSAYTSNCKNGFPGASKFSQATLKUSTVHCQ 690
 Qy 878 YDLPPSSV 885
 |||||
 Db 691 QPIPPRAL 698

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: July 12, 2005, 07:53:17 ; Search time 9805 Seconds
(without alignments)
4482.293 Million cell updates/sec

Title: US-10-751-736-84

Perfect score: 907

Sequence: 1 MDTSLRGVLLSLPVLQLAT.....PAYPTESCHLSVAFVPCPL 907

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 segs, 24227607955 residues

Word size: 1

Total number of hits satisfying chosen parameters: 9408497

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-O=/cpn2_1/USPTO.spool/US10751736/runat_12072005_083640_22171/app_query.fasta_1.1095
-DB=GenEmbl -QFWT=fastap -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-NOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-USER=US10751736 @CGN 1 1 6632 @runat_12072005_083640_22171 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sv.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	2724	6	CQ721366 Sequence
2	907	100.0	2724	6	CQ766844 Sequence
3	907	100.0	2724	6	AR308745 Sequence
4	907	100.0	2724	6	AX549136 Sequence

5	907	100.0	2880	6	AX658241 Sequence
6	907	100.0	2880	9	AF062006 Homo sapi
7	907	100.0	4559	6	BD075815 G protein
8	907	100.0	4570	6	AR447625 Sequence
9	820	90.4	2724	6	AR308754 Sequence
10	730	80.5	2724	9	AF061444 Homo sapi
11	606	66.8	2082	6	BD135244 Novel mam
12	362	39.9	145165	9	AC078860 Homo sapi
13	357	39.4	4095	9	AK075399 Homo sapi
14	202	22.3	606	6	E42915 Novel gonad
15	202	22.3	606	6	AX016185 Sequence
16	71	7.8	162004	9	AC090116 Homo sapi
17	51	5.6	3115	10	AF110818 Mus muscu
18	44	4.9	816	11	BV208561 GPR49_209
19	32	3.5	176092	10	AC129329 Mus muscu
20	32	3.5	195389	2	AC126943 Mus muscu
21	32	3.5	222154	2	AC126127 Rattus no
22	32	3.5	239603	2	AC132780 Rattus no
23	29	3.2	195389	2	AC126943 Mus muscu
24	29	3.2	252477	2	AC136417 Rattus no
25	27	3.0	443	4	AF455789 Sus scro
26	22	2.4	350	5	CR352678 Gallus ga
27	19	2.1	60	6	CQ550152 Sequence
28	18	2.0	284	6	E42919 Novel gonad
29	18	2.0	284	6	AX016194 Sequence
30	18	2.0	1827	6	CQ842816 Sequence
31	18	2.0	1827	9	AK123055 Homo sapi
32	18	2.0	2208	6	AX301830 Sequence
33	18	2.0	2484	9	AF190501 Homo sapi
34	18	2.0	2487	6	AX926297 Sequence
35	18	2.0	2711	6	AX301828 Sequence
36	18	2.0	2901	6	AX301824 Sequence
37	18	2.0	2901	6	AX301833 Sequence
38	18	2.0	3041	6	E42917 Novel gonad
39	18	2.0	3041	6	AX016191 Sequence
40	18	2.0	3119	6	AX549295 Sequence
41	18	2.0	3119	9	AB049405 Homo sapi
42	18	2.0	3273	6	CQ870634 Sequence
43	18	2.0	3286	9	BC047905 Homo sapi
44	18	2.0	3306	6	CQ840811 Sequence
45	18	2.0	3325	9	AY358119 Homo sapi

ALIGNMENTS

RESULT 1	CQ721366	Sequence 7300 from Patent WO02068579.	2724 bp	DNA	linear	PAT 03-FEB-2004
LOCUS	CQ721366					
DEFINITION	CQ721366					
ACCESSION	CQ721366					
VERSION	CQ721366.1	GI:42282223				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.				
AUTHORS		Kits, such as nucleic acid arrays, comprising a majority of				
TITLE		humanexons or transcripts, for detecting expression and other uses				
JOURNAL		thereof				
PATENT		Patent: WO 02068579-A 7300 06-SEP-2002;				
PE CORPORATION		PE Corporation (NY) (US)				
FEATURES		Location/Qualifiers				
SOURCE		1..2724				
		/organism="Homo sapiens"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:9606"				

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2724
Score: 907.00 Matches: 907

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-751-736-84 (1-907) x CQ721366 (1-2724)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
DB 1 ATGGAGCACTCCCGGCTGGTGTCTCTCTTGGCTGTCTGTGAGCTGGCGACC 60

QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
DB 61 GGGGGCAGCTCTCCAGGCTCTGGTGTCTGTGTGGAGGGCTGGCCACACATGTCTATGC 120

QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
DB 121 GAGCCGAGCGGAGGATGTCTCAGGGTGGAGTCTCCGACCTGGGGCTCTCGGAGCTG 180

QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
DB 181 CCTTCCAACTCAGCGTCTTACCTCTCCTACCTAGACCTCAGTATGAACAAATCAGTCAG 240

QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
DB 241 CTGCTCCCGATCCCTCCCGAGTCTCCGCTCTCTGGAGGATTAAGTCTTGGCGGAAAC 300

QY 101 AlaLeuThrTyrIleProGlyGlyAlaPheThrGlyLeuTyrSerLeuValLeuMet 120
DB 301 GCTCTGACATACATTTCCCAAGGGAGCATCTACTGGCCCTTACAGTCTTAAAGTCTTATG 360

QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
DB 361 CTGCAGATAATACGCTAAGACACATGACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420

QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
DB 421 CNAATCCCTGGCTGTGGATGCTAACCAATCAGTATGTGCCCCCAAGCTGTTTCAGTGGC 480

QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
DB 481 CTGCATTCCTTGAGGCACCTGTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540

QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
DB 541 GCTTTTGAAGTTATTCGGCATTCGACGCATGACCTTGGCGCTGAAACAAATATACACCAC 600

QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
DB 601 ATACCAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAAT 660

QY 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
DB 661 AGAATCCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720

QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
DB 721 TTAATTAACAATAACCTTGATGAATTCCTCCACTGCAATTAGGACACTCTCCAACCTTAAA 780

QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
DB 781 GNACTAGAAATTCATAGCAACAAATATCAGTCTGATCTCTGAGAAAGCAATTTGTAGGCAAC 840

QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
DB 841 CCTCTCTTATTAACAATACATTTCTATGACAATCCCATCCAATTTGTTGGGAGATCTGCT 900

QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
DB 901 TTTCAACATTTTACCTGAACCTAAGAACTGACTCTGAATGGTGGCTCTCAAAATTAACGAA 960

QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
DB 961 TTTCTGATTTAACTGGAACTGCAAACTCTGAGAGTCTGACTTTAACTGGAGCACAGATC 1020

DB 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020

QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1021 TCATCTCTCTCTCAAAACCGTCTGCAATCAGTTTACCTATCTCCAAGTCTAGATCTGTCT 1080

QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuLeuGlnLysIleAsp 380
DB 1081 TACAACCTATTAGAAGATTTTACCAGTTTTTTCAGTCTGCCAANAAGCTTCAGANAATTCAC 1140

QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1141 CTAAGACATAATGAAATCTACGAAATTAAGTTCGACACTTTCCAGCAGTGTCTTAGCCTC 1200

QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
DB 1201 CGATCGCTGAAATTTGGCTTGGAAACAAATTTCTATTATTACCCCAATTTTCCACT 1260

QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
DB 1261 TTGCCATCCCCTAATAAAGCTGGACTATGTCGCAACCTCTCTGTCTCTTTTCTATAAAT 1320

QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1321 GGGTTACATGCTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380

QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGT 1440

QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1441 GCATTTGAGTGTGTGAGATGCCCTATAAGATTTCTTAATCAATGGAATTAAGGTGACAA 1500

QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGGAGACCTTTCATAAGAAAGATCTCGAATGTTTCAGGCTCAAGATGAACGT 1560

QY 521 AspLeuGluAspPheLeuLeuAspPheGluAspLeuLysAlaLeuHisSerValGln 540
DB 1561 GACCTTGAAGATTTCTCTGCTTGTGACTTTGAGGAAGACCTGAAAGGCCCTTCATTCAGTGCAG 1620

QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
DB 1621 TGTTCACCTTCCCGAGGCCCTTCAACCTCTGCAACCTGCTGTGATGGCTGGCTGATC 1680

QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATGGAGTGTGGACCATAGCAGTTCGACACTTACTTGTAAATGCTTTGGTGACTTCA 1740

QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACACTGTTAAATGGGGTCTATCGCA 1800

QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 CAGTGAACATGCTCACGGAGTCTCCAGTCCCTGCTGGCTGGTGGATGGCTGTCTACT 1860

QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1861 TTTGGCAGCTTTGACGACATGGTGCCTGGTGGGAGAAATGGGTGGTGGTCCATGTCATT 1920

QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
DB 1921 GGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTTCTCTGCTTACTCTCGCAGCCCTG 1980

QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 1981 GAGCGTGGTGTCTCTGGAATAATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040

QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
DB 2041 CTGAAAGTAAATCAITTTGCTCTGTGCTGCTGCGCTTGACCATGGCGCGCAGTTCCCGCTG 2100

QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCySerLeuProLeuProPheGlyGluPro 720
DB 2101 CTGGTGGGAGCAAGATGCGCTCCCTCTCTGCTGCTGCTTTGGGGAGCC 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuMet 740
DB 2161 AGCACCATGGGTACATGCTGCTCTCATCTGCTCAATTCCTTTGCTTCATGATG 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuLeuLeuLeuLeuLeuLeuLeuLeu 760
DB 2221 ACCATTGGCTACACCAAGCTCTACTCAATTTGGACAGGAGACCTGGAGAATAATTTGG 2280
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 780
DB 2281 GACTGCTCTATGATGATAAACAATGCTGCTCTCTTCCCAACTGCAATCTCTAACTGC 2340
QY 781 ProValAlaPheLeuSerPheSerSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 800
DB 2341 CTGTGGCTTTCTTGTCT 2400
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuLeuLeuLeuLeuLeu 820
DB 2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 821 IleLeuPheAsnProHisPheLysGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 840
DB 2461 ATCTGTTTCAATCTCTCACTTTAAGAGGATCTGCTGCTGCTGCTGCTGCTGCTG 2520
QY 841 TrpThrArgSerLysHisProSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 860
DB 2521 TGGACAAGATCAAAACCAACCAAGCTGATGCTCAATTAATCTGATGATGCTGAA 2580
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerSerSerSerSer 880
DB 2581 TCTGTGACTCAACTCAAGCTTGGTAACTTTACCACTCCAGCTCCAGCATCACTT 2640
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSer 900
DB 2641 CCTCCAGTTCGGTGCATCACCAGCTTATCCAGTACCTGAGTGCAGCTGCCATCT 2700
QY 901 ValAlaPheValProCysLeu 907
DB 2701 GTGGCATTTGTCCCATGCTC 2721

RESULT 2
LOCUS CQ766844 2224-bp DNA linear PAT 07-JUN-2004
DEFINITION Sequence 3 from Patent WO2004005457.
ACCESSION CQ766844
VERSION CQ766844.2 GI:48426391
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Colland, F., Barker, N., Clevers, J.C., Gomez, E., van de Wetering, M.L.
and Suile, E.S.
TITLE THE USE OF SPECIFIED TCF TARGET GENES TO IDENTIFY DRUGS FOR THE
TREATMENT OF CANCER, IN PARTICULAR COLORECTAL CANCER, IN WHICH
TCF/abgr; -CATENIN/WNT SIGNALING PATHWAY IS INVOLVED
JOURNAL Patent: WO 2004005457-A 3 (15-JAN-2005)
COMMENT Kylix B.V. (NL)
FEATURES On Jun 7, 2004 this sequence version replaced gi:44909056.
source Location/Qualifiers
1. .2724
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:

Pred. No.: 0 Length: 2724
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-751-736-84 (1-907) x CQ766844 (1-2724)
QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
DB 1 ATGGACACCTCCCGGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 21 GlyGlySerSerProArgSerGlyValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
DB 61 GGGGGGAGCTCTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 41 GluProAspGlyArgMetLeuLeuLeuValAspCysSerAspLeuGlyValLeuSerGluLeu 60
DB 121 GAGCCCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
DB 181 CCTTCCCACTCAGCGCTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 240
QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuGluLeuLeuLeuLeuLeu 100
DB 241 CTGCTCCCGAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLeuValLeuMet 120
DB 301 GCTCTGACACATACATCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
DB 361 CTGCAAGATAATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTGGAGACCT 420
QY 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
DB 421 CAATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 161 LeuHisSerLeuArgHisLeuThrLeuLeuAspAsnAlaLeuThrGluLeuProValGln 180
DB 481 CTGCATTCCTCAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
DB 541 GTTTTATGAGTTTATCGCATTTGCAAGCCATGACCTTGCCCTGCAACAAATATACACC 600
QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValLeuHisLeuHisAsnAsn 220
DB 601 ATACCAGACTATGCTCTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACA 660
QY 221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
DB 661 AGAATCCACTCCTCGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
DB 721 TTAATTTACAAATAACTTGTGATGATTTCCCATCTGCATTTAGACACTCTCCAACTTAA 780
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
DB 781 GAACTAGGATTTTATAGCAACAATATCAGGTCCATCTGAGAAAGCAATTTGTAGGCAAC 840
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
DB 841 CCTTCTCTTATTAACAATCACTTCTATGCAATCCCATCCCAATTTGTGGGAGATCTGCT 900
QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
DB 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGATGGTGGCTCACAATAACTGAA 960

QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
DB 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1020
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1021 TCATCTCTTCCCTCAACCGCTGCAATCAGTACCTAATCTCCAAGTGTAGATCTGTCT 1080
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1081 TACAACCTATTAGAAAGATTACCCAGTTTTTCAGTCTGCCAAAGCTTCAGAAAAATTGAC 1140
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1141 CTAGACATAATGAATCTACGAAATTAAGTTTGACACTTTCCAGAGTTGCTTAGCCCTC 1200
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
DB 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCTTTTCCACT 1260
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
DB 1261 TTGGCATCCCTAATAAAGCTGACCTATCGTCCAACTCTCGTCTCTTTTCCCTATAACT 1320
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1321 GGGTTACATGGTTTAACTCACCTTAAATTAACAGAAATCATGGCTTTACAGAGCTTGATA 1380
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1381 TCATCTGAAACTTTCCGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTCTGT 1440
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1441 GCATTTGGAGTGTGTGAGAATGCCATATAGATTTCTAATCAATGGAATAAAGGTGACAA 1500
QY 501 SerSerMetAspLeuHisGlyLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGGAGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAGT 1560
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB 1561 GACCTTGAAGATTTCTGTCTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpIleuIle 560
DB 1621 TGTTCACCTTCGCCAGGCCCTTCAAACTCTGAAACCTGTGAAACCTGCTTGATGGCTGTGATC 1680
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAACTTTGTGTGACTTCA 1740
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAATGGGGTCAATGCA 1800
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 GCAGTGAACATGCTACGGGAGTCTCCAGTGGCGTGTGCTGGTGTGTGATCGTTCACT 1860
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1861 TTGGCAGCTTTGGCAGCATGGTGTCTGGTGGGAGATGGGGTGTGTGCCATGTCAAT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
DB 1921 GGTTTTGTCTCCATTTTCTTCCAGATCATCTGTTTTCTTCTTCTTACTCTGCGAGCCCTG 1980
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCAATTTTCTAGC 2040
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700

DB 2041 CTGAAGTAATCATTTTGTCTGTGGCTCTGTGGCTTGACCATGGCCGAGTTCCTCCCTG 2100
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLysLeuProLeuProPheGlyGluPro 720
DB 2101 CTGGTGGCAGCAAGTATGGGCGCTCCCTCTCTGCGTGCCTTTGCCCTTTTGGGGAGCCC 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
DB 2161 AGCACCATGGGTACATGGTGCCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
DB 2221 ACCATTGCTACCAAGCTCTACTGCAATTTGGACAAGGAGACCTTGAGAAATATTGG 2280
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
DB 2281 GACTGCTCTATGGTAAAAACACATTGGCTTGTCTTTCACCAACTGCATCTCAAACTGC 2340
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
DB 2341 CCTGGGCTTCTTGTCTCTTCTCTTATAAACCCTTACATTTATCAGTCTCTGAAGTA 2400
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
DB 2401 ATTAAGTTTATCTTCTGTTGGTAGTCCCACTTCTCTGCATGCTCTCAATCCCTTCTCTAC 2460
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
DB 2461 ATCTGTTCAATCTCTCACTTTAAAGGAGGATCTGGTGAGCTCGAGAAAGCAAACTACGTC 2520
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
DB 2521 TGGACAAGATCAAAACACCAAGCTTGATGCAATTAATCTCTGATGATGTCGAAAAACAG 2580
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
DB 2581 TCCTGTGACTCACTCAAGCCTTGTGAACCTTTACCAGCTCCAGCATCACATTATGACCTG 2640
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
DB 2641 CCTCCAGTTCGGTGCATCACCAGCTTATCCAGTGACTGAGAGTGCATCTTTCTCTCT 2700
QY 901 ValAlaPheValProCysLeu 907
DB 2701 GTGGCATTTGCCCATGTCTC 2721

RESULT 3
AR308745 2724 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 263 from patent US 6555339.
DEFINITION AR308745
ACCESSION AR308745
VERSION AR308745.1 GI:31700274
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 2724)
AUTHORS Liaw,C.W., Behan,D.P. and Chalmers,D.T.
TITLE Non-endogenous, constitutively activated human protein-coupled receptors
JOURNAL Patent: US 6555339-A 263 29-APR-2003;
FEATURES Location/Qualifiers
1..2724
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2724
Score: 507.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0
US-10-751-736-84 (1-907) x AR308745 (1-2724)			
QY	1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr	20	
DB	1 ATGGACACCTCCCGGCTCGGTGTGCTCTCTTCCTGCTGTGCTGCGAGCC	60	
QY	21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysCys	40	
DB	61 GGGGGCAGCTCTCCAGGCTCTGGTGTGCTGTGAGGGGCTGCCACACACTGTCTATG	120	
QY	41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu	60	
DB	121 GAGCCGACGCGGAGATGTCTCAGGGTGAGCTGCTCCGACCTGGGGCTCTCGAGCTG	180	
QY	61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln	80	
DB	181 CCTTCCACCTCAGCGTCTTCACTCTACCTCCTAGACCTCAGTATGAACAACATCAGTCAG	240	
QY	81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn	100	
DB	241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCGAGGAGTACGTCCTTCGGGAAAC	300	
QY	101 AlaLeuThrTyrIleProGlyGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet	120	
DB	301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTACAGTCTTAAAGTCTCTATG	360	
QY	121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu	140	
DB	361 CTGCGAATAATCAGCTAAGACACGTACCCACAGAAAGCTCTCGAATAATTCGGAAGCCTT	420	
QY	141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly	160	
DB	421 CAATCCCTCGTCTGATGCTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480	
QY	161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln	180	
DB	481 CTGCATTCCTCGAGCACCTGTGGCTGGATGACATGCGTTAAACAGAAATCCCGCTCCAG	540	
QY	181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis	200	
DB	541 GCTTTTGAAGTTATCGCATTCGAAGCCATGACCTTGCCCTTGAACAAATAACACCAC	600	
QY	201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValIleLeuHisIleAsnAsn	220	
DB	601 ATACCAGACTATGCTTTTGGAAACCTCTCCAGCTTGTAGTCTTACATCTCCATACCAAT	660	
QY	221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp	240	
DB	661 AGAATCCACTCCTCGGAAAGAAATGCTTTGTATGGCTCCACAGCTAGAGACTTTAGAT	720	
QY	241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys	260	
DB	721 TTTAAATTAACAATAACCTTGATGAATTCCTCCACTGCAATTAGACACTCTCTCAACCTTAA	780	
QY	261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn	280	
DB	781 GAATAGGATTTTCATAGCAACAATATCAGTGCATACCTCGAAGAAGCATTTGTAGGCAAC	840	
QY	281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla	300	
DB	841 CCTTCTCTTATTAACAATACATTTCTATGACATCCATCCATCTTGTGGAGATCTGCT	900	
QY	301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu	320	
DB	901 TTTTCAACATTTACCTGAACATAAGAACACTGACTCTGAATGGTGCCTCACAAAATAACTGAA	960	
QY	321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle	340	
DB	961 TTTTCTGATTTTAACCTGAACTGAAACCTGGAGAGTCTGACTTTTAACCTGGAGACAGATC	1020	

Qy	341	SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer	360
Db	1021	TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAAATCTCCAAGTGTAGATCTGTCT	1080
Qy	361	TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp	380
Db	1081	TACAACCTATTAGAAGATTTACCAGATTTTCAGTCTGCCAAGCTTCAGAAATTTGAC	1140
Qy	381	LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu	400
Db	1141	CTAAGACATAATCAAATCTACGAAATTAAGTTTGACACTTTCCAGCAGTTGCTTAGCCTC	1200
Qy	401	ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr	420
Db	1201	CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTCCACT	1260
Qy	421	LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr	440
Db	1261	TTGCCATCCCTAATAAGCTGGACCTTATGCTCAACTCTCTGTCGTCCTTTTCTCTATACT	1320
Qy	441	GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle	460
Db	1321	GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA	1380
Qy	461	SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys	480
Db	1381	TCATCTGAAACCTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTCTGT	1440
Qy	481	AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn	500
Db	1441	GCATTTGGAGTGTGAGAAATGCTTATAGATTTCTAAATCAATGAATTAAGGTGACAA	1500
Qy	501	SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg	520
Db	1501	AGCAGTATGGACACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560
Qy	521	AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln	540
Db	1561	GACCTTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG	1620
Qy	541	CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle	560
Db	1621	TGTTTACCTTCCCAGGCCCTTCAAAACCTGTGAACACCTGCTTGTATGGCTGGCTGATC	1680
Qy	561	ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer	580
Db	1681	AGAAATTTGAGTGTGGACCATAGCAGTCTGGCAGCTTACTTGTATGCTTTGGTGACTTCA	1740
Qy	581	ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla	600
Db	1741	ACAGTTTTCAGATCCCTCTGTACATTTCCCTCATTAACCTGTTAATTTGGGGTTCATCGCA	1800
Qy	601	AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr	620
Db	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGGCTGCTGGCTGGTGGTGGTGGTGGTGGTGG	1860
Qy	621	PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle	640
Db	1861	TTTTGGCAGCTTTCACGACATAGTGTGCTGGTGGGAGAAATGGGGTGGTGGCATGTCATT	1920
Qy	641	GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu	660
Db	1921	GGTTTTTTGTCCATTTTGTCTTCAAGAAATCATCTGTTTTTCTGTCTTACTCTGGCAGCCCTG	1980
Qy	661	GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer	680
Db	1981	GAGCGTGGGTTCCTCTGTGAAATATTCTGCAAAATTTTGAACCGAAAGCTCCATTTTCTAGC	2040
Qy	681	LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu	700
Db	2041	CTGAAGATTAATCATTTTGTCTGTGCCCTTGTGCCCTTGTGCCCTTGTGCCCTTGTGCCCTT	2100
Qy	701	LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro	720

QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
DB 1021 TCATCTCTTCCCTCAAAACCGTGTGCAATCAGTTACCTAAATCTCAAGTGTAGATCTGTCT 1080
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
DB 1081 TACAACCTATTAGAGATTATCCAGTTTTCNGTCTGCAAAAGCTTCAGAAATTGAC 1140
QY 381 LeuArgHisAsnGluLeuTyrGluLysValAspThrPheGlnGlnLeuLeuSerLeu 400
DB 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCAGCAGTGTGTAGCCTC 1200
QY 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
DB 1201 CGATCGCTGAATTTGGCTTGGACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
QY 421 LeuProSerLeuLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
DB 1261 TTGCCATCCCTAATAAAGCTGGACCTATGTCACACCTCTCTGTCTTTTCTATAACT 1320
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
DB 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
DB 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAAGTGTGT 1440
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
DB 1441 GCATTTGGAGTGTGGAGAAATGCTTATAAGATTCTTAATCAATGGAATAAAGTGACAAAC 1500
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
DB 1501 AGCAGTATGACGACCTTCATAAGAAAGATGCTGGAATCTTTTCAGCTCAAGATGAAGCT 1560
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
DB 1561 GACCTTGAAGATTCTCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
DB 1621 TGTTCACCTTCCCAGGCCCTTCAAAACCTGTGAACACTGCTGTATGGCTGGCTGATC 1680
QY 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
DB 1681 AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACCTACTTGTATGCTTTGGTGACTTCA 1740
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
DB 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCAATTTAAACTGTTAATTTGGGTCATCGCA 1800
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
DB 1801 CAGGTGAACATGCTACGGGAGTCTCCAGTGGCGTGTGGTGGTGGTGGTGGTGGTGGTGGT 1860
QY 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
DB 1861 TTTGGCAGCTTTGGCAGCATGGTGTCTGGTGGGAGATGGGGTGGTGGTGGTGGTGGTGGT 1920
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
DB 1921 GGTTTTTTCTCCATTTTGTCTTCCAGAAATCTGCTTCTGCTTCTGCTTCTGCTGCTGCTG 1980
QY 661 GluArgGlyPheSerValLysTrpSerAlaLysPheGluThrLysAlaProPheSerSer 680
DB 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuLeuThrMetAlaAlaValProLeu 700
DB 2041 CTGAAAGTAAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720

DB 2101 CTGGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGCCCC 2160
QY 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
DB 2161 AGCACCATTGGGTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTTCATGATG 2220
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluLeuIleTrp 760
DB 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGAGAGAAATTTGG 2280
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
DB 2281 GACTGCTCTATGGTAAACACAAATGCTTGTGCTCTTCCCAACTGCATCTCTAAACTGC 2340
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
DB 2341 CTTGTGGCTTCTTCTGCT 2400
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
DB 2401 ATTAAGTTTATCTCTGCTGGTAGTCCCACTTCTGCTGCTCTCTCAATCCCTTCTCTAC 2460
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
DB 2461 ATCTTGTTCATCTCTTAAAGGAGATCTGGTGAGCCTGAGAAAGCAAACTTACGTC 2520
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
DB 2521 TGGACAAGATCAAAACACCAAGCTTGTATGATCAATTAATCTGTATGATGTGCAAAACAG 2580
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
DB 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2640
QY 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
DB 2641 CCTCCAGTTCGTGCCATCACCAGCTTATCCAGTGACTGAGAGTGCCATCTTCTCTCT 2700
QY 901 ValAlaPheValProCysLeu 907
DB 2701 GTGGCAATTTGTCCTCATGCTC 2721

RESULT 5

LOCUS AX658241 2880 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 157 from Patent WO03000928.
ACCESSION AX658241
VERSION AX658241.1 GI:29160776
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Poulsen,H.S., Pedersen,N., Mortensen,S., Sorensen,S.B.,
Petersen,M.W. and Eisner,H.I.
TITLE Methods for identification of cancer cell surface molecules and
cancer specific promoters, and therapeutic uses thereof
JOURNAL Patent:WO 03000928-A 157 03-JAN-2003;
Odin Medical A/S (DK)
FEATURES
Location/Qualifiers
source 1..2880
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 0 Length: 2880
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Db 589 GCTTTTAGAAGTTTATCGCATTCGCAAGCCATGACCTTGGCCCTGAAACAAATAATACACCAC 648
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsn 220
Db 649 ATACAGACTATGCGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTATCATCTCCATACCAAT 708
Qy 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 709 AGAATCCACTCCCTCGGGAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 768
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 769 TTAATATTACAAATAACTGATGAATTCGCCACTGCAATTTAGGACACTCTCCAACTTAAA 828
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 829 GAACTAGGATTTTCATAGCAACAATATCAGGTGCGATACCTGGAAGAGCAATTTGTAGGCAAC 888
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 889 CTTCTCTTATTAACAATACATTTCTATGACAATCCCATCCAAATTTGTTGGGAGATCTGCT 948
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 949 TTTCAACATTTACCTGNACTAAGAACACTGACTCTGAAATGGTGGCTCACAATAAATGAA 1008
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1009 TTTCTCGATTTTAACCTGGAACCTCAAAACCTGGAGAGTCTGACTTTTAACCTGGAGCACAGATC 1068
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1069 TCATCTCTCTCCCAAAACCTCTGCAATCAGTACCTTAATCTCCAAGTGTAGATCTGCT 1128
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1129 TACAACCTATTAGAAGATTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC 1188
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1189 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC 1248
Qy 401 ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1249 CGATCGCTGAAATTTGGCTGGAAACAAATTTGCTATTATTCACCCCAATGCATTTTCCACT 1308
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuLeuSerPheProIleThr 440
Db 1309 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACCTCCTGCTCTTTTCTATTAAT 1368
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1369 GGGTTTACATGGTTAACTCACTTAAATTAACAGGAAATCATGSCCTTACAGAGCTTGATA 1428
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1429 TCATCTGAAATTTTCCAGAACTCAAGGTATAGAAATGCCCTTAGCTTACCAGTCTGT 1488
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1489 GCATTTGGAGTGTGTGAGAAATGCCATATAAGATTTCTAATCAATGAATAAAGGTGACAAC 1548
Qy 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1549 AGCAGTATCGACGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1608
Qy 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1609 GACCTTGAGATTTCTGCTTGACTTTGAGGAAGACCTGAAGGCCCTTCATTCAGTGCAG 1668
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Db 1669 TGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAACACCTGCTTGAATGGCTGGCTGATC 1728

Qy 561 ArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1729 AGAATTTGGAGTGTGACCATAGCAGTTCTGCGACTTACTTGTATATGCTTTGGTGCATTCA 1788
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1789 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAATTTGGGGTCAATCGCA 1848
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1849 GCAGTGAACATGCTCACGGGAGTCTCCAGTCCCGTGGCTGGTGTGGATGCGTTCAT 1908
Qy 621 PheGlySerPheAlaAtrqHisGlyAlaTyrTrpGluAsnGlyValGlyCysHisValIle 640
Db 1909 TTTGGCAGCTTTGACGACATGGTCCCTGGTGGGAGAAATGGGGTTGGTTCGCAATGTCATT 1968
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1969 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTTCTTCTTACTCTGCGAGCCCTG 2028
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2029 GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTTCCATTTTCTAGC 2088
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2089 CTGAAAGTAATCATTTTGTCTGTGCCCTGTGCGCTTGACCATGGCGCAGTTCCCTG 2148
Qy 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2149 CTGGTGGCAGCAAGATATGGCGCTCCCTCTCTGCGCTGCTTTGGGGAGGCC 2208
Qy 721 SerThrMetGlyTyrMetValAlaIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2209 AGCACCATGGGCTACATGGTCCCTCATCTTGTGCTCAATCCCTTGTCTTCTCATGATG 2268
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2269 ACCATTGGCTACACCAAGCTCTACTGCATTTTGGCAAGGGAGACCTGGAGAAATATTGG 2328
Qy 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2329 GACTGCTCTATGGTAAACACACATTCGCCCTGTGCTTCTTCAACAACCTGATCTTAAACTGC 2388
Qy 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2389 CCTGTGGCTTTCTTGTCTCTCTCTTTAATAAACCTTTACATTTATCAGTCTCGAAGTA 2448
Qy 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2449 ATTAAGTTTATCTTCTGGTGGTAGTCCCACTTCTCTGCATGCTCAATCCCTTCTCTAC 2508
Qy 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2509 ATCTTTGTTCAATCTCTCACTTTTAAAGAGGATCTGGTGGCCTGAGAAAGCAAACTACGTC 2568
Qy 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2569 TGGCAAGATCAAAACACCCCAAGCTTGATGCAATTAATCTGATGATGTCGAAAAACAG 2628
Qy 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2629 TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTACAGCTCCAGCATCACATTATGACCTG 2688
Qy 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db 2689 CTTCCAGTTCCGTGCCATCACAGCTTATCCAGTGTATCCAGTGTGAGAGTGCATCTTTCTCT 2748
Qy 901 ValAlaPheValProCysLeu 907
Db 2749 GTGGCATTTTGTCCCATGTCTC 2769

[illegible]

1689	Db		GCATTTGGAGTGTGTGAGATGCCCTATAAGATTTTCTAATCAATGGAATAAAGGTGACAAC	1748
501	Qy	SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg	520	
1749	Db	AGCAGTATGACGACACCTTCTAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1808	
521	Qy	AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln	540	
1809	Db	GACCTTGAAGATTTCTGTGCTTGACTTTGAGGAGACCTGAAAGCCCTTCATTCAAGTCGAG	1868	
541	Qy	CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle	560	
1869	Db	TGTTTCACTTCCCCAGGCCCCCTTCAAAACCCCTGTGAACACCTGCTTGATGCTGGCTGATC	1928	
561	Qy	ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer	580	
1929	Db	AGAAATGGAGTGTGGACCATAGCAGTTCTGTGCATTTACTTGTATGTTTGGTGACTTCA	1988	
581	Qy	ThrValPheArgSerProLeuTyrlIleSerProIleLysLeuLeuIleGlyValIleAla	600	
1989	Db	ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTAAATGGGTCATCGCA	2048	
601	Qy	AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr	620	
2049	Db	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGTGGCTGGTGTGGATGCGTTCACCT	2108	
621	Qy	PheGlySerPheAlaArgHisGlyValATrprpGluAsnGlyValGlyCysHisValIle	640	
2109	Db	TTTGGCAGCTTTGCACACATGGTGCCTGGTGGAGAAATGGGTTGGTGTGCCATGTCAAT	2168	
641	Qy	GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu	660	
2169	Db	GGTTTTTGTCCATTTTTTGCTTCAGATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	2228	
661	Qy	GluArgGlyPheSerValLysTrsSerAlaLysPheGluThrLysAlaProPheSerSer	680	
2229	Db	GAGCGTGGGTTCCTGTGAAATATCTGTCAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2288	
681	Qy	LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu	700	
2289	Db	CTGAAAGTAATCATTTTTGCTCTGTGCCCTGCTGGCCTTGACCAATGGCGCGAGTCCCTG	2348	
701	Qy	LeuGlyGlySerLysTrsGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro	720	
2349	Db	CTGGGTGGCAGCAATATGGGCGCTCCCTCTCTGCTGCTTGGCTTTTGGGAGCCCT	2408	
721	Qy	SerThrMetGlyTrpMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet	740	
2409	Db	AGCACCATGGGCTACATGGTGCCTCTCATCTTGCTCAATTCCTTTCCTCATGATG	2468	
741	Qy	ThrIleAlaTrpThrLysLeuTyrlCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp	760	
2469	Db	ACCATTCCTACACCAAGCTCTACTGCATTTGGACAAGGAGACCTGGAGAAATATTGG	2528	
761	Qy	AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys	780	
2529	Db	GACTGCTCTATGGTAAACAACATTTGCCCTGTGTGCTTCTTCACCACTGCATCCTTAACTGC	2588	
781	Qy	ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal	800	
2589	Db	CCTGTGGCTTCTTGTCTCTCTCTCTTTAATAAACCTTACATTTATCAGTCCCTGAAGTA	2648	
801	Qy	IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyrl	820	
2649	Db	ATTAAAGTTTATCCCTCTCTGGTGTAGTCCCACTTCTGCTGATGTCTCAATCCCTTCTCTAC	2708	
821	Qy	IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrlVal	840	
2709	Db	ATCTTGTTCAATCTCTTAAAGGAGATCTGGTGGCCCTGAGAAAGCAACCTACGTC	2768	
841	Qy	TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln	860	

Db	2769	GGACACAGATCAAAACACCCAGCCTTGATGTCAATTAACTCTGATGATGTCGAAAAACAG	2828
Qy	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880
Db	2829	TCTGTGACTCACTCAGCCTTGGTTAACTTTACAGCTCCAGCATCACTTATGACCTG	2888
Qy	881	ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900
Db	2889	CCTCCAGTTCCTCGTCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT	2948
Qy	901	ValAlaPheValProCysLeu 907	
Db	2949	GTGGCATTTGTCCCATGTCTC 2969	
RESULT 8			
AR447625			
LOCUS AR447625 Sequence 201 from patent US 6673549. linear PAT 20-FEB-2004			
DEFINITION			
ACCESSION AR447625			
VERSION AR447625.1 GI:42675949			
KEYWORDS			
SOURCE Unknown.			
ORGANISM Unknown.			
REFERENCE Unclassified.			
AUTHORS 1 (bases 1 to 4570)			
TITLE Furness,L.M. and Buchbinder,J.L.			
JOURNAL Genes expressed in C3A liver cell cultures treated with steroids			
FEATURES Patent: US 6673549-A 201.06-JAN-2004;			
source Location/Qualifiers			
1. 4570			
/organism="unknown"			
/mol_type="genomic DNA"			
ORIGIN			
Alignment Scores:			
Pred. No.: 0 Length: 4570			
Score: 907.00 Matches: 907			
Percent Similarity: 100.00% Conservative: 0			
Best Local Similarity: 100.00% Mismatches: 0			
Query Match: 100.00% Indels: 0			
DB: 6 Gaps: 0			
US-10-751-736-84 (1-907) x AR447625 (1-4570)			
Qy	1	MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr	20
Db	262	ATGACACACTCCCGGCTCGGTGTCTCTGTCTTGCCTGTGTCTGAGCTGGCGACC	321
Qy	21	GlyCysSerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys	40
Db	322	GGGGGAGCTCTCCAGGTCTGTGTGTGTGTCTGAGGGGCTGCCCCACACACTGTCATTGC	381
Qy	41	GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu	60
Db	382	GAGCCCGACGCGCAGGATGTGCTCAGGGTGGACTGTCTCGACCTGGGGCTCTCGAGCTG	441
Qy	61	ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln	80
Db	442	CCTTCCAACTCAGCGTCTTCACTCTCTACCTAGACCTCAGTATGAAACAACATCAGTCAG	501
Qy	81	LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn	100
Db	502	CTGCTCCGAATCCCTGCCAGTCTCCGCTTCTCGAGGAGTTACGTCTTGGGGAAAC	561
Qy	101	AlaLeuThrTyrIleProIysGlyAlaPheThrGlyLeuTyrSerLeuIysValLeuMet	120
Db	562	GCTCTGACATACATTTCCACAGGGAGCATTCACCTGGGCTTTACAGTCTTAAAGTTCTTATG	621
Qy	121	LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu	140
Db	622	CTGCAAGATATCAGCTTAGACACGTACCCACAGAGAGCTCTGCAGAAATTTGCGAGCCTT	681
Qy	141	GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly	160

Db 682 CAATCCCTGGCTGGATGCTAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 741
QY 161 LeuHisSerLeuArgHisLeuTyrLeuAspAsnAlaLeuThrGluLeuProValGln 180
Db 742 CTGCATTCCTGAGGCACCTGTGGCTGGATGACAATGGGTTAACAGAAATCCCGCTCGAG 801
QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysLeuHis 200
Db 802 GCTTTTAGAAGTTATCGCATTCGCAATTCAGGCATGACCTTGGCCCTGACAAATAACACCC 861
QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Db 862 ATACCAAGATATGCCCTTTGGAACCTCTCAGCTTGGTAGTCTCATCTCCATAACAAT 921
QY 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 922 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 981
QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 982 TTAATTTACAATAACCTTGATGAATTCCTCAGCTGCAATTAGGACACTCTCCAACTTAAA 1041
QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 1042 GAACCTAGGATTTTCATAGCAACAATATCAGTGCATACCTGAGAAAGCATTTGTAGGCAAC 1101
QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 1102 CCTTCTCTTATACATAACATTTCTATGACATCCCATCCATTTGTTGGAGATCTGCT 1161
QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 1162 TTTCAACATTTACTGAACTTAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAA 1221
QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 1222 TTTCTCGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1281
QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1282 TCATCTCTTCTCNAACCTGTGCATTCAGTACCTTAATCTCCAGTGTAGATCTGTCT 1341
QY 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1342 TACAACCTATTAGAAGTTAACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAAAATTGAC 1401
QY 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db 1402 CTAAGACATAATGAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGTCTTAGCCTC 1461
QY 401 ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1462 CGATCGCTGATTTGGCTTGGAAACAATAATGTCTATTATTACCCCAATGCATTTTCCACT 1521
QY 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db 1522 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACCTCTCTGTGCTTTTCTCTATAACT 1581
QY 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1582 GGGTTACATGGTTTAACTCACTTTAAATTTAAACAGGAAATCATGCTTTACAGAGCTTTGATA 1641
QY 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1642 TCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCCCTTATCCACAGTCTGT 1701
QY 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1702 GCATTTGGAGTGTGAGAATGCTATAGATTTCTTAATCAATGGAAATAAGGTGACAC 1761
QY 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520

Db 1762 AGCAGTATGACGACCTTCATTAAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT 1821
QY 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1822 GACCTTTGAAGATTTCTGCTGTTGACTTTGGAGAGACCTCGAAAGCCCTTCAATTCAGTGCAG 1881
QY 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIle 560
Db 1882 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGAACACCTGCTTGAATGGCTGGCTGATC 1941
QY 561 ArgIleGlyValTyrThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1942 AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTATGCTTTGGTGACTTCA 2001
QY 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuLeuGlyValIleAla 600
Db 2002 ACAGTTTTTCAGATCCCTCTGTACATTTCCCAATTTAAACTGTTAATGGGGTCATCGCA 2061
QY 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 2062 GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCTGGTGGTGTGGATGCGTTCACT 2121
QY 621 PheGlySerPheAlaArgHisGlyAlaTyrTyrGluAsnGlyValGlyCysHisValIle 640
Db 2122 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGGAGAAATGGGTTGGTGGCATGTCAAT 2181
QY 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Db 2182 GGTTTTTTGTCCATTTTGTCTTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 2241
QY 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 2242 GAGCGTGGGTTCTCTGTGAAATATTCTGCAAAATTTGAAACACGAAAGCTCCATTTTCTAGC 2301
QY 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2302 CTGAAGATTAATCATTTTGTCTCTGTGCCCTGTGCTGGCCTTGACCATGGCCGAGTTCCCTCG 2361
QY 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Db 2362 CTGGGTGGCAGCAGATATGGCGCTCCCTCTCTGCTGCTTGTGCTTTTGGGAGCC 2421
QY 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2422 AGCACCATGGCTTACATGCTGCTCATCTTGTCTCAATTCCTTGTCTCTCATGATG 2481
QY 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr 760
Db 2482 ACCATGGCTTACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTGG 2541
QY 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db 2542 GACTGCTCTATGTTAAACACATTTGCCCTGTGCTTCTTCCCACTGCATCTCTAACTGC 2601
QY 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db 2602 CCGTGGCTTCTTGTGCTTCTCTCTCTTAAATAAACCTTATACATTTATCAGTCTCGAAGTA 2661
QY 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 2662 ATTAAGTTTATCTTCTGTTGGTAGTCCCACTTCTGTCATGCTCAATCCCCCTTCTCTAC 2721
QY 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db 2722 ATCTTGTTCATCTCTCACCTTTAAGGAGGATCTGGTGAGCTTGAGAAAGCAACCTACGTC 2781
QY 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db 2782 TGGACAAGATCAAAACACCAAGCTTGATGTCAATTAATCTCTGTGATGATCGAAAAACAG 2841
QY 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db 2842 TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTTACAGCTCCAGCATCATCTTATGACCTG 2901

Db 1561 GACCTTGAAGATTTCCTGCTTGAAGACCTTGAAGACCTTGAAGACCTTCAATTCAGTGCAG 1620
 Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuile 560
 Db 1621 TGTTCACCTTCCCGAGCCCTTCAAAACCTGTGAACACCTGCTTGCATGGCTGCTGATC 1680
 Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
 Db 1681 AGAATTGGAGTGGACCATAGCAGTCTGGCACCTTACTTGTAAATCTTGTGTGACTTCA 1740
 Qy 581 ThrValPheArgSerProLeuTrpIleSerProIleLysLeuLeuIleGlyValIleAla 600
 Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCATTAACCTGTTAAATGGGGTCATCGCA 1800
 Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
 Db 1801 GCAGTGAACATGCTACGGGAGTCTCCAGTGGCGTCTGGCTGGTGGTGGTGGTGGTGGT 1860
 Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
 Db 1861 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGT 1920
 Qy 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
 Db 1921 GGTITTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGAGCCCTG 1980
 Qy 661 GluArgGlyPheSerValLysTrpSerAlaLysPheGluThrLysAlaProPheSerSer 680
 Db 1981 GAGCGTGGTCTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
 Qy 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
 Db 2041 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
 Qy 701 LeuGlyGlySerLysTrpGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
 Db 2101 CTGGGTGGCAGCAGATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
 Qy 721 SerThrMetGlyTrpMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 740
 Db 2161 AGCACCATGGGTACATGCTGCTCATCTGCTCATCTGCTCATCTGCTGCTGCTGCTGCTG 2220
 Qy 741 ThrIleAlaTrpThrLysLeuTrpCysAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 760
 Db 2221 ACCATTGCTTACACCAAGCTTACTGCAATTTGGACAGGAGACCTGGAGAATAATTTGG 2280
 Qy 761 AspCysSerMetVal-LysHisIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 780
 Db 2281 GACTGCTCTATGAA-AAAAACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2339
 Qy 780 sProValAlaPheLeuSerPheSerSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 800
 Db 2340 CCTGTGGCTTCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTG 2399
 Qy 800 lIleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTrp 820
 Db 2400 AATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2459
 Qy 820 rIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTrpVa 840
 Db 2460 CATCTTGTTCATCTCTCACTTTAAGGAGGATCTGGTGGAGCCTGGAGAACAACCTAC 2519
 Qy 840 lTrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
 Db 2520 CTGGACAGATCAAAACCCAGCTTGAATGCAATTAATTAATTAATTAATTAATTAATTAAT 2579
 Qy 860 nSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLe 880
 Db 2580 GTCTGTGACTCAACTCAAGCCTTGGTAACCTTTACAGCTCCAGCATCACTATGACCT 2639
 Qy 880 uProProSerSerValProSerProAlaTrpProValThrGluSerCysHisLeuSerSe 900
 Db 2640 GCCTCCAGTTCGTGCCATCACCAGCTTATCAGTGAAGTGGAGAGTGCATCTTTCTTCYC 2699

QY 900 rValalaPheValProCysLeu 907
 Db 2700 TGTGGCATTTGTCCCATGTCTC 2721
 RESULT 10
 AF061444
 LOCUS
 DEFINITION
 Homo sapiens G protein-coupled receptor LGR5 (LGR5) mRNA, complete cds.
 ACCESSION AF061444
 VERSION AF061444.1 GI:3985471
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 2724)
 AUTHORS Hsu,S.Y., Liang,S.G. and Haueh,A.J.
 TITLE Characterization of two LGR genes homologous to gonadotropin and a thyrotropin receptors with extracellular leucine-rich repeats and a G protein-coupled, seven-transmembrane region
 JOURNAL Mol. Endocrinol. 12 (12), 1830-1845 (1998)
 MEDLINE 99055210
 PUBMED 9849958
 REFERENCE 2 (bases 1 to 2724)
 AUTHORS Hsu,S.Y., Liang,S.G. and Haueh,A.J.W.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-1998) Gyn/Ob, Stanford University, 300 Pasteur Dr., Stanford, CA 95305-5317, USA
 FEATURES
 source
 1..2724
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 1..2724
 /gene="LGR5"
 1..2724
 /gene="LGR5"
 /note="Leucine-rich repeat containing G protein-coupled receptor; similar to glycoprotein hormone receptors; similar to FSH receptor, LH receptor and TSH receptor"
 /codon_start=1
 /product="G protein-coupled receptor LGR5"
 /protein_id="AAC77911.1"
 /db_xref="GI:3985472"
 /translation="MDTSRLGVLSLPLVLQATGSSPSRGLVLRGCPTHCHCEPDG RMLLVDCSDLSLSELPNSLVFTSYDLDMNNISQLLPNPLPSLHFLEELAGNAL TYIPKAGFTGLYSLKVLMLQNNQRLHVPTALQNLRSLSQSLRDANDHISYVPPSCFSG LHSRLWDDNALTEI PQAFPSLSALQAMTALNKIHHI PDYAFGNLSVSVVHLH NNRHSLGKCFDGLHSLDTLDNYNLDLDEFTAITRLSNLKEGLFHSNNIRSIPEKA FVGNPLIT IHFYDNP IQFVGRSA FOHLPELRTLTLNGASQITERPDLTGNTANLESLT LTGAQISLPQTVCNQLNQLVLDLSYNLLEDLPFSVCQKLOKIDLRHNEIYEIKVD TFQQLSRLSNLNAIKAIHPNAPSTLPSLKLDSLNSLSSFPITGLHGLHLKL TGNHALQSLISENPELVEMPYAYQCAFVGCENAYKISQNNKGNDSMDDLHK KDAGNFQAGDRLEDLDEEDLKALHSVQCSPSPGPFKCEHLLDGWLIRIGVYT IAVLALTCAVLTSTVFRSPLYISPKLLIGVIAAVNMLTGSSAVLAGVDAFTFGSF ARHGAWENGVCVHIGFLSIPASESSVFLTLAALERGSVKYSAKPRTKAPFSSLK VIIICALLALTMVAVPLLGSKYKASPLCLPLPFGEPSTMGVMYVALLILNSLCFLMM TIATYTKLVCNLDKGLNIWDCSWKHFAHLLFTNCLNCPVAFISFSSLSLNTFTFSP EYIKFLLVVPPLPACLNPLLYLFLNFHKFDLVSRLKQTYVTWTSKSHSLMSINSD VEQKSCDSTQALVTFTSSITVDLPSSVPSYPVPTESCHLSSVAFVPL"
 ORIGIN
 Alignment Scores:
 Pred. No.: 0 Length: 2724
 Score: 730.00 Matches: 905
 Percent Similarity: 99.56% Conservative: 0
 Best Local Similarity: 99.56% Mismatches: 2
 Query Match: 80.49% Indels: 4
 DB: 9 Gaps: 0
 US-10-751-736-84 (1-907) x AF061444 (1-2724)

Qy	1	Met	Asp	Thr	Ser	Arg	Leu	Gly	Val	Leu	Ser	Leu	Pro	Val	Leu	Leu	Gln	Leu	Ala	Thr	20
Db	1	ATG	A	C	A	C	T	C	C	G	G	T	G	C	T	C	T	G	C	T	60
Qy	21	Gly	Gly	Ser	Ser	Pro	Arg	Ser	Gly	Val	Leu	Leu	Arg	Gly	Cys	Pro	Thr	His	Cys	His	40
Db	61	GGG	G	G	G	C	A	G	T	C	T	C	C	A	G	T	C	T	G	T	120
Qy	41	Glu	Pro	Asp	Gly	Arg	Met	Leu	Leu	Arg	Val	Asp	Cys	Ser	Asp	Leu	Gly	Leu	Ser	Glu	60
Db	121	GAG	C	C	C	G	A	G	G	A	T	T	G	C	T	C	A	G	G	T	180
Qy	61	Pro	Ser	Asn	Leu	Ser	Val	Phe	Thr	Ser	Tyr	Leu	Asp	Leu	Ser	Met	Asn	Asn	Leu	Ser	80
Db	181	CGT	T	C	C	A	A	C	T	C	A	C	T	C	T	C	A	G	A	C	240
Qy	81	Leu	Leu	Pro	Asn	Pro	Leu	Pro	Ser	Leu	Arg	Phe	Leu	Glu	Glu	Leu	Arg	Leu	Ala	Gly	100
Db	241	CTG	T	C	C	G	A	T	C	C	C	T	C	C	A	T	C	C	A		299
Qy	100	Ala	Ala	Leu	Thr	Tyr	Leu	Pro	Lys	Gly	Ala	Phe	Thr	Gly	Leu	Tyr	Ser	Leu	Lys	Val	120
Db	300	CGC	T	C	T	G	A	C	A	T	A	T	C	C	A	G	G	A	G	A	359
Qy	120	T	Leu	Gln	Asn	Asn	Gln	Leu	Arg	His	Val	Pro	Thr	Glu	Ala	Leu	Gln	Asn	Leu	Arg	140
Db	360	GCT	G	A	G	A	T	A	T	A	T	C	A	G	T	A	C	C	A		419
Qy	140	u	Gln	Ser	Leu	Arg	Leu	Asp	Ala	Asn	His	Leu	Ser	Tyr	Val	Pro	Ser	Cys	Phe	Ser	160
Db	420	TCA	T	C	C	T	C	G	G	T	C	T	G	A	T	G	T	A	C		479
Qy	160	Y	Leu	His	Ser	Leu	Arg	His	Leu	Tyr	Leu	Asp	Asp	Asn	Ala	Leu	Thr	Glu	Leu	Pro	180
Db	480	CCT	G	C	A	T	T	C	C	T	G	A	G	C	A	C	T	G	G	A	539
Qy	180	n	Ala	Phe	Arg	Ser	Leu	Ser	Ala	Leu	Gln	Ala	Met	Thr	Leu	Ala	Leu	Asn	Lys	Leu	200
Db	540	GGC	T	T	T	A	G	A	G	T	T	A	T	C	G	G	A	T	G		599
Qy	200	s	ile	Pro	Asp	Tyr	Ala	Phe	Gly	Asn	Leu	Ser	Ser	Leu	-	Val	Val	Leu	His	Leu	220
Db	600	CAT	ACC	A	G	A	C	T	A	T	G	C	T	C	C	A	G	-	CTG		658
Qy	220	s	n	Arg	Ile	His	Ser	Leu	Gly	Lys	Lys	Cys	Phe	Asp	Gly	Leu	His	Ser	Leu	Thr	240
Db	659	ATA	G	A	A	T	CCA	T	CC	T	C	G	G	A	AA	A	A	T	G		718
Qy	240	s	p	Leu	Asn	Tyr	Asn	Asn	Leu	Asp	Glu	Phe	Pro	Thr	Ala	Ile	Arg	Thr	Leu	Ser	260
Db	719	ATT	T	AA	A	AT	T	CA	A	T	T	CA	A	T	T	CC	C	A	T		778
Qy	260	y	s	Glu	Leu	Gly	Phe	His	Ser	Asn	Asn	Ile	Arg	Ser	Ile	Pro	Glu	Leu	Ala	Phe	280
Db	779	A	AG	AA	C	T	AG	G	A	T	T	CA	T	AG	T	C	A	T	C		838
Qy	280	s	n	Pro	Ser	Leu	Ile	Thr	Ile	His	Phe	Tyr	Asp	Asn	Pro	Ile	Gln	Phe	Val	Gly	300
Db	839	ACC	CT	T	C	T	CT	T	ATT	A	CA	A	T	CA	T	T	AT	G	A		8

Qy	360	erTyrAsnLeuLeuGluuAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleA	380
Db	1079	CTTACAACCTATTAGAAAGATTACCACGATTTTTCAGTCTGCCAAAAGCITTCAGAAAATTG	1138
Qy	380	spLeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerL	400
Db	1139	ACCTAAGACACATAATAGAAATCTACGAATAATAAGTTGACACTTTCACAGCAGTGTCTTAGCC	1198
Qy	400	euArgSerLeuAsnLeuAlaTIPAsnLysIleAlaIleIleHisProAsnAlaPheSerT	420
Db	1199	TCCGATCGCTGAATTTGGCTTGGAAACAAATATGCTATTATTACACCCCAATGCATTTTCCA	1258
Qy	420	hrLeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleT	440
Db	1259	CTTTGGCCATCCCTAATAAAGCTGACCTATCGTCCAACCTCCTGTGCTCTTTTCCCTATAA	1318
Qy	440	hrGlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuI	460
Db	1319	CTGGGTTTACATGGTTTAACTCACTTAAATTAACAGGAAATCAATGCTTACACAGCTTGA	1378
Qy	460	leSerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysC	480
Db	1379	TATCATCTGAAACCTTTCCAGACTCAAGGTTATAGAAATGCCATTATGCTTACACGTGCT	1438
Qy	480	ysAlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspA	500
Db	1439	GTGCAATTTGGAGTGTGTGAGAAATGCCCTATAAGATTTCTAATCAATGAATAAAGGTGACA	1498
Qy	500	snSerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluA	520
Db	1499	ACAGCAGTATGGACGACCTTCATAAGAAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAAC	1558
Qy	520	rgAspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValG	540
Db	1559	GTGACCTTGAGAGATTTCCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGC	1618
Qy	540	lnCysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleuI	560
Db	1619	AGTCTTCACCTTCCCGACGCCCTTCAACCCCTGTGAACACCTGCTTGATGGCTGGCTGA	1678
Qy	560	leArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrS	580
Db	1679	TCAGAAATTTGGAGTGTGGACCATAGCAGTCTTGGCAGCTTACTTGTAAATGCTTTTGGTGACTT	1738
Qy	580	erThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleA	600
Db	1739	CAAAGATTTTCAGATCCCTCTGTACATTTTCCCCCATTAACGTGTTAATTTGGGGTCATCG	1798
Qy	600	laAlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheT	620
Db	1799	CAGCAGTGAACATGCTCACGGGAGTCTCCAGTCCGCTGGCTGGTGGTGGATGCGTTCA	1858
Qy	620	hrPheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValI	640
Db	1859	CTTTTGGCAGCTTTTGCACGACATGCTGCCCTGGTGGGAGATGGGGTTGGTGTGCCATGTCA	1918
Qy	640	leGlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaL	660
Db	1919	TTGGTTTTTGTGTCATTTTGTCTTCAGAAATCATCTGTTTTTCCCTGCTTACTCTTGCAGCCC	1978
Qy	660	euGluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerS	680
Db	1979	TGGAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTA	2038
Qy	680	erLeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProL	700
Db	2039	GCCTGAAAGTAATCATTTTGTCTGTGCCCTGTGCTGGCTTTGACCATGGCGCAGTTTCCCC	2098
Qy	700	euLeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluP	720
Db	2099	TGCTGGGGTGGCAGCAGATGGCGCCCTCCCTCTCTGCTGCCCTTTGCCCTTTTGGGGAGC	2158
Qy	720	roSerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetM	740


```

Db      2159  CGAGCACCAGGCGCTACATGGTGGCTCTCATCTGCTCAATTCCTTGGCTTCTCTCATGA 2218
QY      740  etThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleT 760
Db      2219  TGACCATTCCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAAATTT 2278
QY      760  rPaspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnC 780
Db      2279  GGGACTGCTCTATGGTAAACACATTCCTGCTGCTTTCACCAACTGCATCTAAACT 2338
QY      780  yProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluV 800
Db      2339  GCCCTGGCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2398
QY      800  alIleLysPheIleLeuLeuValValProLeuProAlaCysLeuLeuAsnProLeuLeuT 820
Db      2399  TAAATTAAGTTTATCTCTCTGTTGGTAGTCCCACTTCTGCAATGCTCAATCCCTCTCT 2458
QY      820  yIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrV 840
Db      2459  ACATCTTGTTCATCTCACTTTAAGGAGGATCTGGTGGCTTGAGAAAGCAACCTAGC 2518
QY      840  alTnThrArgSerLysHisProSerLeuMetSerIleLeuSerAspAspValGluLysG 860
Db      2519  TCTGGCAAGATCAAAACACCCCAAGCTTGTATGTCATTAATTAACCTGTGATGATGTCG 2578
QY      860  InSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspL 880
Db      2579  AGTCTGTGACTCACTCAAGCTTGTGTAACCTTTACCAAGCTCCAGCATCACTTAGACC 2638
QY      880  euProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerS 900
Db      2639  TGCCTCCCAAGTTCGGTGCCATCACCACTTATCCAGTACTGAGAGTGCATCTTCTCT 2698
QY      900  erValAlaPheValProCysLeu 907
Db      2699  CTGTGGCATTTGTCCCATGCTCTC 2721

RESULT 11
LOCUS   BD135244
DEFINITION Novel mammalian G protein-coupled receptor having extracellular
ACCESSION BD135244
VERSION   BD135244.1 GI:23230189
KEYWORDS JP 2002507406-A/2.
SOURCE   Homo sapiens
ORGANISM Homo sapiens
REFERENCE Hsueh,A.J.W., Hsu,S.Y., Liang,S.G. and Spe,P.J.V.D.
AUTHORS 1 (bases 1 to 2082)
TITLE    Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
JOURNAL  Hsueh,A.J.W., Hsu,S.Y., Liang,S.G. and Spe,P.J.V.D.
        Novel mammalian G protein-coupled receptor having extracellular
        leucine-rich repeating domain
        Patent: JP 2002507406-A 2 12-MAR-2002;
        THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY,
        AKZO NOBEL NV
COMMENT  OS Homo sapiens (human)
        PN JP 2002507406-A/2
        PD 12-MAR-2002
        PF 25-MAR-1999 JP 2000537903
        PR 26-MAR-1998 US 60/079501
        PI ARON J W HSUEH, SHEAU YU HSU, SHAN GUANG LIANG, PETRUS JOHANNES
        P1 VAN DER SPEK
        PC C12N15/09,A01K67/027,C07K14/705,C07K16/28,C12N1/15,C12N1/19,
        PC C12N1/21,
        PC C12N5/10,C12P21/08,C12N15/00,C12N5/00
        CC Novel mammalian G protein-coupled receptor having CC
        extracellular
        CC leucine-rich repeating domain
        FH Key
        FT Location/Qualifiers
        FT source 1..2082

```

```

FT      Location/Qualifiers
FEATURES Source
Pred. No.: 0 Length: 2082
Score: 606.00 Matches: 692
Percent Similarity: 99.71% Conservative: 0
Best Local Similarity: 99.71% Mismatches: 1
Query Match: 66.81% Indels: 2
DB: Gaps: 0

ORIGIN
US-10-751-736-84 (1-907) x BD135244 (1-2082)
QY      215  LeuHisLeuHisAsnAsnArgIleHisSerLeuGlyLysCysPheAspGlyLeuHis 234
Db      1   CTACATCTCCATAACAATAGATCCACTCCCTGGGAAGAAATGCTTTGATGGCTCCAC 60
QY      235  SerLeuGluThrLeuAspLeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArg 254
Db      61  AGCCTAGAGACTTTAGATTTAAATTACAAATACCTTGATGATGATTTCCCACTGCAATTAGG 120
QY      255  ThrLeuSerAsnLeuLysGluLeuGlyPheHisSerAsnAsnIleArgSerIleProGlu 274
Db      121  ACATCTCCCAACTTAAAGGAACCTAGGATTTCTAGCAACAATATCAGGTCGATACCTGAG 180
QY      275  LysAlaPheValGlyAsnProSerLeuIleThrIleHisPheTyrAspAsnProIleGln 294
Db      181  AAAGCATTTTGTAGGCAACCTCTCTTATTACAAATACATTTCTATGACAAATCCCATCCAA 240
QY      295  PheValGlyArgSerAlaPheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGly 314
Db      241  TTGTGGGAGATCTGCTTTTCAACATTTACCTGAACTAAGAACTGACTCTGATGGT 300
QY      315  AlaSerGlnIleThrGluPheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThr 334
Db      301  GCCTCACAATAAATGAAATTTCTGATTTAACTGGAACCTGCAACCTGGAGAGCTGTGACT 360
QY      335  LeuThrGlyAlaGlnIleSerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeu 354
Db      361  TTAACCTGGAGACAGATCTCATCTCTCTCAACCGCTCTGCAATCAGTTACCTAATCTC 420
QY      355  GlnValLeuAspLeuSerTyrAsnLeuLeuGluAspLeuProSerPheSerValCysGln 374
Db      421  CAAGTGTAGATCTGTCTTACAACTATTAGAAAGATTTACCCAGTTTTCAGTCTGCCAA 480
QY      375  LysLeuGlnLysIleAspLeuArgHisAsnGluIleTyrGluIleLysValAspThrPhe 394
Db      481  AGCTTCAGAAATTTGACCTTAGACATAATGAAATCTACGAAATTAAGATTGACACTTTC 540
QY      395  GlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAlaThrAsnLysIleAlaIleHis 414
Db      541  CAGCAGTTGCTTAGCTCGATCGCTGAAATTTGGCTTGGAAACAAATTTGCTATTATTAC 600
QY      415  ProAsnAlaPheSerThrLeuProSerLeuIleLysLeuAspLeuSerSerLeuLeu 434
Db      601  CCCAATGCAATTTTCCACTTTGCCATCCCTAAATAAGCTGGACCTATCGTCCAACTCCCTG 660
QY      435  SerSerPheProIleThrGlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHis 454
Db      661  TGTCTTTTCTTAACTGGGTTCATGTTTAACTCCTTAAATTAACAGGAATCAT 720
QY      455  AlaLeuGlnSerLeu-IleSerSerGluAsnPheProGluLeuLysValIleGluMetPr 474
Db      721  GCCTTACAGAG-CTGGATATCATCTGAAACTTTCCAGAACTTCCAGAACTCAAGGTNATGAAATGCC 779
QY      474  cTyrAlaTyrGlnCysCysAlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGln 494
Db      780  TTATGCTTACCAGTGTCTGTGCATTTTGGAGTGTGTGAGAAATGCCATAAGATTCTTAATCA 839

```


QY 494 nTrpAsnLysGlyAspAsnSerSerMetAspAspLeuHisLysLysAspAlaGlyMetPh 514
DB 840 ATGGAAATAAGGTGACAAAGAGAGATATGGACACCTTCATAGAAGAGATGCTGGAAATGTT 899
QY 514 eGlnAlaGlnAspGluArgAspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLys 534
DB 900 TCAGGCTCAAGATGAAGTGAACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACTGNA 959
QY 534 sAlaLeuHisSerValGlnCysSerProGlyProPheLysProCysGluHisLe 554
DB 960 AGCCCTTCATTCAAGTCAAGTGTACCTTCCCGAGGCCCTTCAAAACCCCTGTGAACACT 1019
QY 554 uLeuAspGlyTrpLeuLeuArgLysValTrpThrLeuAlaValLeuAlaLeuThrCys 574
DB 1020 GCTTCATGCTGGCTGCATCAGAAATTTGGAGTGTGGACCATAGCAGTTCCTGGCACTTACTTGT 1079
QY 574 sAsnAlaLeuValThrSerThrValPheArgSerProLeuTyrIleSerProIleLysLe 594
DB 1080 TAATGCTTTGGTGAATTCACAGTTCACAGTTCCTGATCCCTCTGTACATTTCCCCCAATTAAC 1139
QY 594 uLeuLeuGlyValIleAlaAlaValAsnMetLeuThrGlyValSerSerAlaValLeuAl 614
DB 1140 GTTAATTTGGGGTCATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGC 1199
QY 614 aGlyValAlaAspAlaPheThrPheGlySerPheAlaArgHisGlyAlaTTrpGluAsnG 634
DB 1200 TGGTGTGAATGGCTTTCACATTTTGGGAGCTTTTGCAGCAGCATGCTGCTGGTGGGAAGTGG 1259
QY 634 yValGlyCysHisValIleGlyPheLeuSerIlePheAlaSerGluSerSerValPheLe 654
DB 1260 GGTGGTGGCAATGATGCTTTTGGTTCATTTTGGTTCATTTTGGTTCAGAAATCATCTGTTTCT 1319
QY 654 uLeuThrLeuAlaAlaLeuGluArgGlyPheSerValLysTyrSerAlaLysPheGluTh 674
DB 1320 GCTTACTCTGGCAGCCCTGGAGCGTGGGTCTCTGTGAATAATTTCTGCGCCCTGCTGGCTTGAAC 1379
QY 674 rLysAlaProPheSerSerLeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuTh 694
DB 1380 GAAAGCTTCATTTCTAGCCTGAAGATGAATCAATTTTGTCTGTGCGCTGCTGGCTTGAAC 1439
QY 694 rMetAlaAlaValProLeuLeuGlyLysSerLysTyrGlyAlaSerProLeuCysLeuPr 714
DB 1440 CATGGCCGAGTTCCTGCTGGTGGGAGCAGATATGGCGCTCCCTCTCTGCTGCTGCC 1499
QY 714 oLeuProPheGlyGluProSerThrMetGlyTyrMetValAlaLeuLeuLeuAsnSe 734
DB 1500 TTTGGCTTTTGGGAGGCCAGCACCATGGGCTACATGGTGGCTCTCATCTTCTCAATTC 1559
QY 734 rLeuCysPheLeuMetMetThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysG 754
DB 1560 CCTTTGCTTCTCATGATGACCATTTGCTTACCAAGCTCTACTGCAATTTGGAAGGG 1619
QY 754 yAspLeuGluAsnIleTrpAspCysSerMetValLysHisIleAlaLeuLeuLeuPheTh 774
DB 1620 AGACCTGGAGATATTTGGGACTGCTCATGTGTAAACACATTTGGCTGCTGCTTCTCAC 1679
QY 774 rAnCysIleLeuAsnCysProValAlaPheLeuSerPheSerSerLeuIleAsnLeuTh 794
DB 1680 CAACCTGCATCCTAAACCTGCTGGCTTCTTGTCTTCTCTCTCTTCTTAAATAAACCTTAC 1739
QY 794 rPheIleSerProGluValIleLysPheIleLeuLeuValValProLeuProAlaCys 814
DB 1740 ATTTATCATGCTCTGAAGTAATTAAGTTTATCTTCTGGTGGTAGTCCCACTTCTCGCATG 1799
QY 814 sLeuAsnProLeuLeuTyrIleLeuPheAsnProHisPheLysGluAspLeuValSerLe 834
DB 1800 TCTCAATCCCTTCTCATCATCTTGTTCATCTCTTAAAGGAGGATCTGGTGGCT 1859
QY 834 uArgLysGlnThrTyrValTrpThrArgSerLysHisProSerLeuMetSerIleAsnSe 854
DB 1860 GAGAAAGCAACACTAGTCTGGCAGAGATCAAAACACCAAGCTTGTATCAATTAATTC 1919

QY 854 rAspAspValGluLysGlnSerCysAspSerThrGlnAlaLeuValThrPheThrSerSe 874
DB 1920 TGATGATGTCGAAAAACAGTCTGTGACTCACTCAAGCTTGGTAACTTTACGAGCTC 1979
QY 874 rSerIleThrTyrAspLeuProProSerSerValProSerProAlaTyrProValThrGl 894
DB 1980 CAGCATCACTTATGACTTCCCTCCAGTTCCTGGTGCATCACCAGCTTATCCAGTCACTGA 2039
QY 894 uSerCysHisLysSerSerValAlaPheValProCysLeu 907
DB 2040 GAGCTGCCATCTTCTCTGTGGCATTTGTCCCATGTCTC 2079
RESULT 12
AC078860/c
LOCUS
DEFINITION Homo sapiens 12 BAC RP11-186F10 (Roswell Park Cancer Institute
Human BAC Library) complete sequence.
AC078860
VERSION AC078860.19 GI:13491193
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 145165)
AUTHORS Murzyn,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooke,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbacia,J.,
Benton,J., Blmage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,J., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogues,M., Hollway,C., Hollins,B., Homsif.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsone,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Louisged,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,N., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S.,
Ogih,M., Okunou,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojehokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shooshkar,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williamson,A.,
Wlaczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y.F., Zhou,J.,
Zorrilla,S., Kucherlapati,R. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 145165)
Worley,K.C.
Direct Submission
Submitted (07-AUG-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (31-MAR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (01-APR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (15-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 145165)
Worley, K.C.
Direct Submission
Submitted (23-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 31, 2001 this sequence version replaced gi:13324671.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 145169
Phrap values in estimate: 144455
Average error rate (BCM-Phrap estimate): 5.28339e-05
Fraction of Phrap values less than 40 : 0.0137759
Number of consensus changing edits: 19
Number of N's in consensus : 0

----- Consensus changing edits -----

Position	Original+Context	Edited+Context
6882	agatgtgat(a)gctgtggtt	agatgtgat(g)gctgtggtt
12828	tttgggtcca(n)tcgaacctcc	tttgggtcca(c)tcgaacctcc
13151	aaacaacat(n)caatagataa	aaacaacat(a)caatagataa
28957	ttgtattgca(n)atggcatctt	ttgtattgca(a)atggcatctt
37139	atcgtattg(n)tgtaatttt	atcgtattg(t)tgtaatttt
58030	tttttatt(n)tatagttaa	tttttatt(t)tatagttaa
58290	tcacgacct(a)tcagtccct	tcacgacct(a)tcagtccct
64729	ctatctctac(n)aaagaaaga	ctatctctac(g)aaagaaaga
92472	aatggcttt(g)aatgggccc	aatggcttt(g)aatgggccc
92662	cttagactgt(n)aatgcaggga	cttagactgt(g)aatgcaggga
103773	gctaattttt(n)catctttagt	gctaattttt(a)catctttagt
110456	cacatgcatt(n)ggattttgtc	cacatgcatt(a)ggattttgtc
117534	tcaagcttgt(n)atttcaggtt	tcaagcttgt(a)atttcaggtt
117581	cagatcatct(n)agggaatgct	cagatcatct(g)agggaatgct
125485	gcttacacta(n)aaatgaacc	gcttacacta(g)aaatgaacc
133723	agcatgggtgt(n)nncacttaat	agcatgggtgt(t)ggcacttaat
133724	gcattggtgt(n)ncacttaata	gcattggtgt(g)gcacttaata
133725	catggtgttn(n)cacttaataa	catggtgttg(g)cacttaataa
140801	aaatatgtgc(n)taataaagt	aaatatgtgc(c)taataaagt

----- Distribution of Quality < 40 Bases -----

	5	10	15	20	25	30	35	40
1000								
900								*
800								*
700								*
600								*
500								*
400	*						*	*
300	*	*				*	*	*
200	*	*	*			*	*	*
100	*	*	*	*		*	*	*
0	*	*	*	*	*	*	*	*
Phrap Value Range	5	10	15	20	25	30	35	40

FEATURES

Source

Version: 1.01 qxf.

Location/Qualifiers

1..145165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosomes="12"
/clone="RP11-186F10"
1..2004
/note="overlaps bases 52259..54262 of clone AC073612"
/function="clone overlap"
complement(6..223)
/rpt_family="LIME1"
complement(404..698)
/rpt_family="LIM4"
complement(833..1077)
/rpt_family="LIM4"
complement(1228..1482)
/rpt_family="LIM4"
complement(1740..2021)
/rpt_family="ALUJb"
complement(2098..2151)
/rpt_family="LIMEC"
complement(2152..2449)

misc_feature

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

Alignment Scores:

Pred. No.: 0

Score: 362.00

Percent Similarity: 100.00%

Length: 145165

Matches: 362

Conservative: 0


```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.91% Indels: 0
DB: 9 Gaps: 0

US-10-751-736-84 (1-907) x AC078860 (1-145165)

QY 546 GlyProPheLeuProCysGluHisLeuLeuAapGlyTrpLeuileAArgileGlyValTrp 565
DB 72537 GGCCTTCCTCAAAACCTGTGAACACCTGCTTTGATGGCTGCTGATCAGAAATGGAGTGG 72478

QY 566 ThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSer 585
DB 72477 ACCATAGCAGTTCTGGCATTACTTGTAAATGCTTTGTGGTACCTCAACAGTTTTCAGATCC 72418

QY 586 ProLeuTyrlleSerProIleLysLeuLeuileGlyValIleAlaAlaValAsnMetLeu 605
DB 72417 CCTCTGTACATTTCCCTCCCATTAACCTGTTAAATTTGGGGTCATCGCAGCAGTGAACATGCTC 72358

QY 606 ThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAla 625
DB 72357 ACGGAGATCTCCAGTGCCTGCTGCTGTGGATGCGTTTACATTTTGGCAGCTTTCGA 72298

QY 626 ArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIle 645
DB 72297 CGACATGGTCTGCTGGGAGAAATGGGGTGGTGGCCATGTCATTGTTTGTCCATT 72238

QY 646 PheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSer 665
DB 72237 TTTGCTTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCTCGAGCGTGGGTTCCT 72178

QY 666 ValLysTyrlleSerAlaValPheGluThrLysAlaProPheSerSerLeuLysValIleIle 685
DB 72177 GTGAATATTTCTGCAAAATTTGAAACGAAAGTCCCATTTTCTAGCCTGAAAGTAATCAT 72118

QY 686 LeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLys 705
DB 72117 TTGCTCTGGCCCTGCTGGCCTTGACCATGGCCGAGTTCCTGCTGGTGGCAGCAAG 72058

QY 706 TyrGlyAlaSerProLeuLeuProLeuProPheGlyGluProSerThrMetGlyTyr 725
DB 72057 TATGGCGCTCCCTCTGCTGCTGCTTTGGCTTTTGGGAGCCACCACTGGGCTAC 71998

QY 726 MetValAlaLeuLeuLeuLeuAsnSerLeuLysPheLeuMetMetThrIleAlaTyrlle 745
DB 71997 ATGGTTCGCTCTCATCTTGCTCAATTTCCCTTTGCTTCTCATGATGACCAATGCCCTACAC 71938

QY 746 LysLeuTyrlleCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetVal 765
DB 71937 AAGCTCTACTGCAATTTGGACAAGGGAGACCTGGAGAAATATTTGGGACTGCTCTATGTA 71878

QY 766 LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeu 785
DB 71877 AAACACATTTGCCCTGTGCTTCTACCAACTGCATCTAACTGACCTGCTGCTTCTTG 71818

QY 786 SerPheSerSerLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeu 805
DB 71817 TCCTTCTCTCTTTAATAAACCTTACATTTATCAGTCTCGAAGTAATTAAGTTTATCCTT 71758

QY 806 LeuValValProLeuProAlaCysLeuAsnProLeuLeuTyrlleLeuPheAsnPro 825
DB 71757 CTGGTGGTAGTCCCACTTCTGCTGCTCATGCTCAATCCCTTCTCTACATCTTGTTCATCT 71698

QY 826 HisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrlleThrArgSerLys 845
DB 71697 CACTTTAAGAGGAGATCTGGTAGCCTCGAGAAAGCAAACTACCTCTGGCAAGATCAAAA 71638

QY 846 HisProSerLeuMetSerIleAsnSerAspValGluLysGlnSerCysAspSerThr 865
DB 71637 CACCAAGCTTGTATGCAATTAATCTGTATGATGTGCAAAACAGTCTCTGTGACTCAACT 71578

QY 866 GlnAlaLeuValThrPheThrSerSerSerIleThrTyrlleAspLeuProProSerVal 885
DB 71577 CAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCACTTATGACCTGCTCCCAAGTTCCGCTG 71518
```

```
QY 886 ProSerProAlaTyrlleProValThrGluSerCysHisLeuSerSerValAlaPheValPro 905
DB 71517 CCATCACCAGCTTATCCAGTGCAGTGCAGAGTGCCTCTCTCTGTGGCATTTGTCCCA 71458

QY 906 CysLeu 907
DB 71457 TGTCTC 71452

RESULT 13
LOCUS AK075399
DEFINITION Homo sapiens cDNA PSEC089 fis, clone PLACE1001148, highly similar
to Homo sapiens orphan G protein-coupled receptor HG38 mRNA.
ACCESSION AK075399
VERSION AK075399.1 GI:22761463
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Kawai-Hio,Y., Hayashi,K.,
Ishii,S., Saito,K., Yamamoto,J., Wakamatsu,A., Nagai,T.,
Nakamura,Y., Sato,H., Nagahari,K., Sugano,S. and Isogai,T.
TITLE HRI human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4095)
AUTHORS Isogai,T. and Yamamoto,J.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Takao Isogai, Helix Research Institute,
Genomics Laboratory; 1532-3 Yana Kisarazu, Chiba 292-0812, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT HRI human cDNA sequencing project; cDNA 5'- and 3'-end one pass
sequencing, clone selection and full insert sequencing; Helix
Research Institute (supported by Japan Key Technology Center etc.);
cDNA library construction; Institute of Medical Science, University
of Tokyo, Laboratory of Genome Structure, Human Genome Center.
FEATURES
source
1..4095
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1001148"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="cloning vector: pME18SFL3"
1168..1451
/note="This 284 bases segment is not present in AF062006
and AF061444, probably caused by alternative splicing"
misc_feature
ORIGIN
Alignment Scores:
Pred. No.: 0 Length: 4095
Score: 357.00 Matches: 357
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 39.36% Indels: 0
Gaps: 0
US-10-751-736-84 (1-907) x AK075399 (1-4095)

QY 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
DB 98 ATGGACACCTCCCGCTCGGTGTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 157

QY 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
DB 158 GGGGGCAGCTCTCCCGAGTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCTATGC 217

QY 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
DB 218 GAGCCCGCAGCGCAGGATGTTGTCAGGGTGGAGTCTGCTCGAGCTGCGAGCTGGGGCTCTCGGAGCTG 277
```


QY 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnLeuSerGln 80
 DB 278 CTTTCCAACTCAGCGCTTCCCTCCTAGACCTCAGTATGAACAACATCAGTCAG 337
 QY 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
 DB 338 CTGCTCCCAATCCCTGCCAGCTCCCGCTTCTGGAGAGGTACGTCTTGGCGGAAC 397
 QY 101 AlaLeuThrTyrLeuProGlyAlaPheThrGlyLeuTyrSerLeuLeuValLeuMet 120
 DB 398 GCTCTGACATACATCCCAAGGGAGCATTCACCTGGCTTTACAGTCTTAAAGTCTTATG 457
 QY 121 LeuGlnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
 DB 458 CTGCAGATAATCAGCTAAGACACGTAACCCACAGAAGCTCTGCAGAAATTTGGCAAGCCTT 517
 QY 141 GlnSerLeuArgLeuAspAlaLeuHisIleSerTyrValProSerCysPheSerGly 160
 DB 518 CAATCCCTGGCTGGATGCTAACCACTAGCTATGTGCCCAAGCTGTTTCAGTGGC 577
 QY 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
 DB 578 CTGCATTCCTGAGGACCTGTGGCTGGATGACATGCGTTAACAGAAATCCCGCTCCAG 637
 QY 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
 DB 638 GCTTTTAGAAGTTATCGGCAATGCAAGCATGACCTTGCCCTGAAACAAATACACCAC 697
 QY 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisIleHisAsnAn 220
 DB 698 ATACCAAGATATGCTCTTTGGAAACCTCTCCAGCTGGTAGTCTACATCTCCATAACAAT 757
 QY 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
 DB 758 AGAATCCATCTCCCTGGGAAGAAATGCTTTGATGGCTCCACAGCTTAGAGCTTAGAT 817
 QY 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
 DB 818 TTAAATACAAATAACCTTGATGAATCCCACTGCAATTAGGACACTCTCCAACCTTAA 877
 QY 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
 DB 878 GAACATAGATTTTCATAGCAACAATATCAGGTGATACCTGAGAAAGCATTTGTAGGCAAC 937
 QY 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
 DB 938 CTTTCTCTTATACATACATCTTATGACAAATCCCATCCATTTGTTGGAGATCTGCT 997
 QY 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
 DB 998 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTCAATGGTGCCTCACAAATACTGAA 1057
 QY 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
 DB 1058 TTTCTCTGATTTAATCGAACTGCAACCTAGAGAGTCTGACTTTAATCGGAGCACAGATC 1117
 QY 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeu 357
 DB 1118 TCATCTCTTCTCAAAACCGTCTGCAATCATGTTACCTTAATCTCCAGTGGCTG 1168
 RESULT 14
 E42915
 LOCUS E42915 606 bp DNA linear PAT 31-JAN-2002
 DEFINITION Novel gonadotropin receptor.
 ACCESSION E42915
 VERSION E42915.1 GI:18627551
 KEYWORDS JP 2000125884-A/3.
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 606)
 AUTHORS Heikopu, Y.C. and Supekku, P.J.V.D.

TITLE
JOURNAL

Novel gonadotropin receptor
 Patent: JP 2000125884-A 3 09-MAY-2000;
 AKZO NOBEL NV

COMMENT

OS Unidentified
 PN JP 2000125884-A/3
 PD 09-MAY-2000
 PF 05-FEB-1999 JP 1999029120
 PF 06-FEB-1998 EP 98200357.6, 27-JUL-1998 EP 98202519.9, PR
 24-SEP-1998 EP 98203213.8
 PI YUDEISU CHRISTINA HEIKOPU, PETORIYUSU JOHANNES VAN DEL SUPEKKU
 PC C12N15/09, A61K38/00, A61P15/08, A61P43/00, C07K14/72, C12N1/15, PC
 C12N1/19
 PC C12N1/21, C12N5/10, C12P21/02, C12Q1/68, G01N33/15, G01N33/50, PC
 G01N33/566,
 PC C12N15/00, A61K37/02, C12N5/00
 CC Strandedness: Single;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..606 /organism='Unidentified'.

FEATURES

source

Location/Qualifiers
 1..606
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.: 7,04e-196 Length: 606
 Score: 202.00 Matches: 202
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 22.27% Indels: 0
 DB: 6 Gaps: 0

US-10-751-736-84 (1-907) x E42915 (1-606)

QY 562 IleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThr 581
 DB 1 ATTGAGTGTGGACCATAGCAGTCTTGCCACTTACTTGTAAATGCTTTGGTGCATTCACCA 60
 QY 582 ValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAla 601
 DB 61 GTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATGGGGTTCATCGCAGCA 120
 QY 602 ValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPhe 621
 DB 121 GTGAACATGCTCAGGGAGTCTCCAGTCCGCTGGCTGGTGGATGGGTTCATCTTTT 180
 QY 622 GlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGly 641
 DB 181 GGCAGCTTTGCACGACATGGTGCCTGGTGGGAGATGGGGTGGTGGCATGTCTATTGGT 240
 QY 642 PheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeuGlu 661
 DB 241 TTTTGTCCATTTTGTCTTTCAGAAATCATCTGTTTCTCTCTACTCTGCGAGCCCTGGAG 300
 QY 662 ArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeu 681
 DB 301 CGTGGGTCTCTGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTG 360
 QY 682 LysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeu 701
 DB 361 AAAATAATCATTTGTCTGTGCCCTGTGCTGGCTTGGACCATGGCCGAGTTCCTCTGCTG 420
 QY 702 GlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSer 721
 DB 421 GGTGGCAGCAGATATGGGCGCTCCCTCTCTGCTGCTCTTGGCTTTTGGGGAGCCAGC 480
 QY 722 ThrMetGlyTyrMetValAlaLeuLeuLeuAsnSerLeuCysPheLeuMetThr 741
 DB 481 ACCATGGGCTACATGCTGCTCTCATCTTGTCAATTCCTTTGCTTCTCATGATGACC 540

Qy	742	IleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAenlleTrrpAsp	761
Db	541	ATTCCCTACACCAAGCTCTACTGCAATTTCGACAAGGAGACCTGGAGAATATTTCGGAC	600
Qy	762	CysSer	763
Db	601	 TGCTCT	606
RESULT 15			
AX016185	AX016185	Sequence 3 from Patent EP0950711.	
ACCESSION	AX016185		
VERSION	AX016185.1	GI:10041794	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE	1		
AUTHORS	Heikoop,J.C. and Van Der Spek,P.J.		
TITLE	Gonadotropin receptor		
JOURNAL	Patent: EP 0950711-A 3 20-OCT-1999;		
AKZO NOBEL NV (NL)			
FEATURES	Location/Qualifiers		
source	1..606		
/organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"			
ORIGIN			
Alignment Scores:			
Pred. No.:	7.04e-196	Length:	606
Score:	202.00	Matches:	202
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	22.27%	Indels:	0
DB:	6	Gaps:	0
US-10-751-736-84 (1-907) x AX016185 (1-606)			
Qy	562	IleGlyValTrrpThrIleAlaValLeuAlaLeuThrCysAenAlaLeuValThrSerThr	581
Db	1	ATTGGAGTGGCAGCATAGCAGTCTTGCGACTTACTTGTAAATCTTTGGTGACTTCAACA	60
Qy	582	ValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAla	601
Db	61	GTTTTCAGATCCCTCTGTACATTTCCCCCATTAACATGTTAAATGGGGTCATCGCAGCA	120
Qy	602	ValAenMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPhe	621
Db	121	GTGAACATGCTCAGGGAGTCTCCAGTGCCTGTGGCTGTGTGGATGGCTTCACTTTT	180
Qy	622	GlySerPheAlaArgHisGlyValATrrpGluAenGlyValGlyCysHisValIleGly	641
Db	181	GGCAGCTTTGCACACATGGTGCCCTCTGGTGGAGAAATGGGGTGTGTGCCATGTCTTGGT	240
Qy	642	PheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGlu	661
Db	241	TTTTTGTCATTTTGTCTTCAGATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTGGAG	300
Qy	662	ArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeu	681
Db	301	CGTGGGTTCTCTGTGAATAATTCGCAAAATTTGAACGAAAGCTCCATTTTCTTAGCCTG	360
Qy	682	LysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeu	701
Db	361	AAAGTAATCATTTTGCTCTGTGGCCCTGTGCGCTTGACATGGCGCAGTTCCCTGCTGCTG	420
Qy	702	GlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSer	721
Db	421	GGTGGCAGCAAGTATGGGCGCTCCCTCTCTGCTGCTTGGCTTTTGGGAGCCGACG	480

Db 61 GGGGGCAGCTCTCCAGGCTCGGTGTGTCTGAGGGCTGCCACACACTGTCTATGTC 120
Qy 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Db 121 GAGCCCGAGCGGAGAGATGTCTCAGGGTGGACTGCTCCGACTGGGGCTCTCGAGCTG 180
Qy 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Db 181 CTTTCCACCTCAGGGCTTCACTCTCTACCTAGACCTCAGTATGAAACACATCACTCAG 240
Qy 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuLeuLeuAlaGlyAsn 100
Db 241 CTGCTCCCGAATCCCTCGCCAGTCTCCGCTTCTCGAGGAGTACGCTCTTGGCGGAAAC 300
Qy 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Db 301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCCCTTACAGTCTTTAAAGTCTTATG 360
Qy 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Db 361 CTGCAGATATATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTGGCAGCCTT 420
Qy 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Db 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Qy 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
Db 481 CTGCATTCCTCGAGCACCTGTGGCTGGATGCAATGCGTTAAACAGAAATCCCGCTCCAG 540
Qy 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Db 541 GCTTTTAGAAGTTATCGGCATTCAGCCATGACCTTGGCCCTCGAACAAATACACCAC 600
Qy 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValLeuHisLeuHisAsnAsn 220
Db 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Qy 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db 661 AGAATCCCATCCCTCGGGAAGAAATGCTTTGTATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db 721 TTAATTTACATPACCTTGATGAATTTCCCACTGCATTTAGAGCACTCTCCAACTTAAA 780
Qy 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db 781 GAACCTAGGATTTTCATAGCAACAATATCAGGTGATACCTGAGAAAGCATTTGTAGGCAAC 840
Qy 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db 841 CCTTCTCTTATTAACAATAATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Qy 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db 901 TTTCAACATTTACCTGGAACATAAGAACACTGACTCTGAATGGTGGCTTCACAAATAACTGAA 960
Qy 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db 961 TTTCTGATTTAACTGGAACCTGCAACCTGGAGAGTCTGACTTTAACTGGACACAGATC 1020
Qy 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db 1021 TCATCTCTTCTCAAAACCTCTGCAATCAGTTACCTTAATCTCCAAAGTGTAGATCTGCT 1080
Qy 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db 1081 TACAACTTATTAGAAGATTTACCCAGTTTTTTCAGTCTGCAAAAGCTTCAGAAATTTGAC 1140
Qy 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400

Db 1141 CTAAGACATAATAGAAATCTACGAAATTAAGTTGACACTTTTCCAGCAGTTGCTTAGCCTC 1200
Qy 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCATTTTCCACT 1260
Qy 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuLeuSerSerPheProIleThr 440
Db 1261 TTGCCATCCCTAAATAAGCTGGACCTATGCTCCAACCTCTCTGTCGTCTTTTCTATAACT 1320
Qy 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGAAATCATGCTTTACAGAGCTTGATA 1380
Qy 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db 1381 TCACTGAAACTTTTCCAGAACTCAGGTTATAGAAATGCCTTATGCTTACCAGTGTGT 1440
Qy 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db 1441 GCATTTTGAGTGTGTGAGAAATGCTTCAATGAATTTCTAATCAATGAATAAAGGTGACAAC 1500
Qy 501 SerSerMetAspAspLeuHisLysIlyshaspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db 1501 AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Qy 521 AspLeuGluAspPheLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db 1561 GACCTTGAAAGATTTCTGCTTGCATTTGAGGAGNACCTGAAAGCCCTTCATTCAGTGGCAG 1620
Qy 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleLeu 560
Db 1621 TGTTCACCTTCCCAGGCCCCCTTCAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC 1680
Qy 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db 1681 AGAATTGAGTGTGGACCATGACGATTCCTGACACTTACTTGTAAATGCTTTGGTGAATTC 1740
Qy 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAATTTGGGGTCAATCGCA 1800
Qy 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db 1801 GCAGTAGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGCTGGTGGATGCGTCACT 1860
Qy 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db 1861 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAAT 1920
Qy 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db 1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db 1981 GAGCGTGGTCTCTGTGTAATATTTCTGCAAAATTTTGAACCGAAAGCTCCATTTTCTAGC 2040
Qy 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db 2041 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGGCTGGGAGAAATGGGGTTGGTTGCCATGTCAAT 2100
Qy 701 LeuGlyGlySerLysTyrGlyValSerProLeuCysLeuProLeuPheGlyGluPro 720
Db 2101 CTGGGTGGCAGCAAGATATGGCGCTCCCTCTCTGCTGCTGCTTTGCCCTTTTGGGGAGCCC 2160
Qy 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db 2161 AGCACCATGGGCTACATGTCGCTCTCATCTTGCTCAATTCCTTTGCTTCCTCATGATG 2220
Qy 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db 2221 ACCATTGCTTACACCAGCTCTACTGCAATTTTGGCAAGGGGAGACCTGTGAGAATATTTGG 2280

Qy	761	AspCysSerMetValValHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys	780
Db	2281	GACTGCCTCTATGGTAAACACATTTGCCCTGTGTCTTCCACCACTGCATCTCAAACTGC	2340
Qy	781	ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal	800
Db	2341	CCGTGGCTTCTTGTCTCTCTCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA	2400
Qy	801	IleIlePheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr	820
Db	2401	ATTAAGTTTATCCTTCTGGTGGTAGTCCCACTTCTCTGCATGTCTCAATCCCTTCTCTAC	2460
Qy	821	IleIlePheAsnProHisPheIleYsGluAspLeuValSerLeuArgIleYsGlnThrTyrVal	840
Db	2461	ATCTGTGTCAATCTCTCACTTTAAGAGGAGTCTGTGTAGCCTCGAAGAACCACTACGTC	2520
Qy	841	TrpThrArgSerIleYsProSerLeuMetSerIleAsnSerAspAspValGluIleYsGln	860
Db	2521	TGGACAAGATCAAAACACCCCAAGCTTCATGTCAATTAACTCTGATGATGTCGAAAAACAG	2580
Qy	861	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerSerIleThrTyrAspLeu	880
Db	2581	TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACCAGCTCCAGCATCATTTATGACCTG	2640
Qy	881	ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer	900
Db	2641	CCCTCCAGTTCCGTGCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT	2700
Qy	901	ValAlaPheValProCysLeu	907
Db	2701	GTGGCATTTGTGCCATGTCTC	2721

RESULT 2

US-09-976-594-201
; Sequence 201, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 201
; LENGTH: 4570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 205542.2
; NAME/KEY: unsure
; LOCATION: 3900, 3919, 3934
; OTHER INFORMATION: a, t, c, g, or other
; US-09-976-594-201

50 Column

```

Alignment Scores:
Pred. No.: 0 Length: 4570
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-976-594-201 (1-4570)

Qy 1 MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr 20
Db 262 ATGGACACTCCCGGCTCGGTGTGCTCTGCTTCCTGCTGCTGCTGCTGCTGCGGAC 321

```


41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
121 GAGCCGACGCGGAGGATGTGCTCAGGGTGAGCTGCTCCGACCTCGGGCTCTCGGAGCTG 180
61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
181 CTTTCCAACTCAGCGCTCTTACCTCTCCCTAGACCTCAGTATGAACAACATCAGTCAG 240
81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuAlaGlyAsn 100
241 CTGCTCCCGAATCCCTCGCCAGTCTCCGCTTCTCGAGGAGTACGCTTCGCGGAAAC 300
101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
301 GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGGCGCTTTACAGTCTTAAAGTCTTATG 360
121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
361 CTGCAGAAATAATCAGCTAAGACACGTACCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
421 CAATCCCTCGCTCTGGATGCTTAACACATCAGTATGTGCCCCCAAGCTGTTTCAGTGGC 480
161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
481 CTGCATTCCTCGAGCACCTGTGGTGGATGACAATGCGTTAAACAGAAATCCCGCTCCAG 540
181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
541 GCTTTTGAAGTTTATCGCATGTCAGCCATGCAAGCCATGACCTTGGCCCTGAAACAAATACACCAC 600
201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGTAGTCTTACATCTCCATAAACAT 660
221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
661 AGAATCCACTCCTCGGAGAAAGAAATGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
721 TTAATAATCAATAACCTTGATGAATTCCTCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
781 GAACTAGGATTTATAGCAACAAATATCAGTGCATACCTGAGAAAGCAATTTGTAGGCAAC 840
281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
841 CCTTCTCTATTACAAATACATTTCTATGACAATCCCATCCCAATTTGTTGGAGATCTGCT 900
301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
901 TTTCAACATTTACCTGGAACATAAGACACTGACTCTGAATGTGGCTCAACAAATACTGAA 960
321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
961 TTTCTCTGATTTTACCTGGAACTGCAAACTGGAGAGTCTGACTTTTAACCTGGAGCACAGATC 1020
341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
1021 TCATCTCTCTCAAAACGCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
1081 TACAACTCTATTAGAAAGATTTACCAAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
1141 CTAAGACATAATGAAATCTACGAAATTAAGTTTGACACTTTTCCAGCAGTGTGCTTAGCTC 1200

401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
1201 CGATCGCTGAAATTTGGCTTGGAAACAAAATTTGCTATTATTATTCACCCCAATGCATTTTCCACT 1260
421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACCTCTCTGTCTCTTTCTCTTAAT 1320
441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTCTTACAGAGCTTGATA 1380
461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAAGTGTCT 1440
481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
1441 GCATTTGGAGTGTGTGAGAAATGCTTAAAGATTCTTAATCAATGAATANAAGGTGACAAAC 1500
501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
1501 AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaIleHisSerValGln 540
1561 GACCTTGAAGATTTCTCTGCTTACCTTTGAGGAAGACCTGAAAGCCCTTTCATTCAGTGCAG 1620
541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
1621 TGTTCACCTTCCCAAGGCCCTTCAACCTGTGTGAACACCTCTGTGAGGCTGGCTGATC 1680
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
1681 AGAATGGAGTGTGGACCATAGCAGTCTGGCACCTTACTTGTAAATGCTTTGTGTGACTTCA 1740
581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAATTTGGGGTCACTCGCA 1800
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
1801 GAGTGAACATGCTCACGGAGTCTCCAGTCCGCTGGCTGGCTGGTGGATCGGTTCAC 1860
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
1861 TTTGGCACCTTTGCACGACATGCTGCTGGTGGGAGATGGGTTGGTGGTTCATTCAT 1920
641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrIleAlaLeu 660
1921 GGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
1981 GAGCGTGGTGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
2041 CTGAAAGTAATCAATTTTGTCTGTGCCCTGTGGCTTGGACCATGGCCGAGTTCCTCCCTG 2100
701 LeuGlyGlySerLysTyrGlyAlaSerProLeuLysLeuProLeuProPheGlyGluPro 720
2101 CTGGTGGCAGCAAGATATGGCGCTCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGCC 2160
721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
2161 AGCACCATGGGTACATGCTCATCTGTCTCAATTCCTTTGCTTCTCTCAATG 2220
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
2221 ACCATGTCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTCGAGAAATATTTGG 2280
761 AspCysSerMetVal-LysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCy 780

CITY:	VALLEY FORGE
STATE:	PA
COUNTRY:	USA
ZIP:	19482
COMPUTER READABLE FORM:	
MEDIUM TYPE:	Diskette
OPERATING SYSTEM:	DOS
SOFTWARE:	FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:	
APPLICATION NUMBER:	US/08/866,757
FILING DATE:	30-MAY-1997
CLASSIFICATION:	536
PRIOR APPLICATION DATA:	
APPLICATION NUMBER:	
FILING DATE:	
ATTORNEY/AGENT INFORMATION:	
NAME:	PRESTIA, PAUL F.
REGISTRATION NUMBER:	23,031
REFERENCE/DOCKET NUMBER:	GH-70055
TELECOMMUNICATION INFORMATION:	
TELEPHONE:	610-407-0700
TELEFAX:	610-407-0701
TELEX:	846169
INFORMATION FOR SEQ ID NO:	1:
SEQUENCE CHARACTERISTICS:	
LENGTH:	4203 base pairs
TYPE:	nucleic acid
STRANDEDNESS:	single
TOPOLOGY:	linear
MOLECULE TYPE:	CDNA
US-08-866-757-1	
Alignment Scores:	
Pred. No.:	0.031 Length: 4203
Score:	13.00 Matches: 13
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	1.43% Indels: 0
DB:	2 Gaps: 0
US-10-751-736-84 (1-907) x US-08-866-757-1 (1-4203)	
Qy 208 AsnLeuSerSerLeuValValLeuHisLeuHisAsnAen 220	
Db 158 AACCTTTCAAGCGTGGTAGTGTCATCTTCATAACAAT 196	
RESULT 8	
US-09-153-593-1	
; Sequence 1, Application US/09153593A	
; Patent No. 6174994	
GENERAL INFORMATION:	
; APPLICANT: ELSHOURBAGY, NABIL A	
; APPLICANT: LI, XIAOTONG	
; APPLICANT: BERGSMÄ, DERK J	
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)	
; FILE REFERENCE: GH-70055-1	
; CURRENT APPLICATION NUMBER: US/09/153,593A	
; CURRENT FILING DATE: 1998-09-15	
; EARLIER APPLICATION NUMBER: 08/866,757	
; EARLIER FILING DATE: 1997-05-30	
; NUMBER OF SEQ ID NOS: 4	
; SOFTWARE: FastSeq for Windows Version 3.0	
; SEQ ID NO 1	
; LENGTH: 4203	
; TYPE: DNA	
; ORGANISM: HOMO SAPIENS	
US-09-153-593-1	
Alignment Scores:	
Pred. No.:	0.031 Length: 4203
Score:	13.00 Matches: 13
Percent Similarity:	100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 3 Gaps: 0

US-10-751-736-84 (1-907) x US-09-153-593-1 (1-4203)

Qy 208 AsnLeuSerLeuValValLeuHisLeuHisAsnAsn 220
Db 158 AACCTTTCAGCCCTGGTAGTCTGCACTTTCATACAAAT 196

RESULT 9

US-09-397-787-240
; Sequence 240, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; TITLE OF INVENTION: CANCER THERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 240
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-397-787-240

Alignment Scores:
Pred. No.: 0.283 Length: 453
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 3 Gaps: 0

US-10-751-736-84 (1-907) x US-09-397-787-240 (1-453)

Qy 810 ProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db 120 CCCCTGCTGCTGCTGCTCAACCCACTGCTGTAC 152

RESULT 10

US-09-556-916-25
; Sequence 25, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 25
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-556-916-25

Alignment Scores:
Pred. No.: 128 Length: 2019
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-556-916-25 (1-2019)

Qy 708 AlaSerProLeuCysLeuProLeuPro 716
Db 1892 GCCTCCCTTTGTGCTGCTTTGCCA 1918

RESULT 11

US-09-556-916-27
; Sequence 27, Application US/09556916
; Patent No. 6548271
; GENERAL INFORMATION:
; APPLICANT: Turner, Alex
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Nehls, Michael
; APPLICANT: Friedrich, Glenn
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. 6548271el Human Transporter Proteins
; FILE REFERENCE: 8535-041-999
; CURRENT APPLICATION NUMBER: US/09/556,916
; CURRENT FILING DATE: 2000-04-21
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 2019
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-556-916-27

Alignment Scores:
Pred. No.: 128 Length: 2019
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-556-916-27 (1-2019)

Qy 708 AlaSerProLeuCysLeuProLeuPro 716
Db 1892 GCCTCCCTTTGTGCTGCTTTGCCA 1918

RESULT 12

US-09-949-016-5657
; Sequence 5657, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5657
; LENGTH: 2680
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5657

Alignment Scores:
Pred. No.: 172 Length: 2680
Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-949-016-5657 (1-2680)

QY 95 LeuArgLeuAlaGlyAsnAlaLeuThr 103

DB 2324 CTGAGACTGGCTGGAATGCTCTGACT 2350

RESULT 13

US-09-799-451-422

; Sequence 422, Application US/09799451

; Patent No. 6783969

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle

; APPLICANT: Asundi, Vinod

; APPLICANT: Ren, Feiyan

; APPLICANT: Zhang, Jie

; APPLICANT: Xue, Aidong J.

; APPLICANT: Zhao, Qing A.

; APPLICANT: Wang, Jian-Rui

; APPLICANT: Ma, Yungqing

; APPLICANT: Yamazaki, Victoria

; APPLICANT: Chen, Rui-hong

; APPLICANT: Wang, Zhiwei

; APPLICANT: Wang, Dunrui

; APPLICANT: Yang, Yonghong

; APPLICANT: Wehrman, Tom

; APPLICANT: Ghosh, Reena

; APPLICANT: Drmanac, Radoje T.

; TITLE OF INVENTION: No. 6783969el Nucleic Acids and

; FILE REFERENCE: 803

; CURRENT APPLICATION NUMBER: US/09/799,451

; CURRENT FILING DATE: 2001-03-05

; NUMBER OF SEQ ID NOS: 948

; SOFTWARE: pt_FL_genes Version 2.0

; SEQ ID NO 422

; LENGTH: 2820

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (265)..(2577)

US-09-799-451-422

Alignment Scores:

Pred. No.: 181 Length: 2820
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-09-799-451-422 (1-2820)

QY 708 AlaSerProLeuCysLeuProLeuPro 716

DB 2156 GCCTCCCTTGTGCTGCTGCTTGGCA 2182

RESULT 14

US-10-101-464A-458

; Sequence 458, Application US/10101464A

; Patent No. 6768041

; GENERAL INFORMATION:

; APPLICANT: Strabala, Timothy

; APPLICANT: Nieuwenhuizen, Nicolaas

; APPLICANT: Higgins, Colleen M.

; TITLE OF INVENTION: Compositions Isolated from Plant Cells

; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020c2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; CURRENT FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 458
; LENGTH: 3600
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-458

Alignment Scores:

Pred. No.: 234 Length: 3600
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-84 (1-907) x US-10-101-464A-458 (1-3600)

QY 354 LeuGlnValLeuAspLeuSerTyrAsn 362

DB 598 CTTGAGTTCGACCTGCTCTTATAAT 624

RESULT 15

US-07-792-865D-1/c

; Sequence 1, Application US/07792865D

; Patent No. 5646247

; GENERAL INFORMATION:

; APPLICANT: John W. Barnwell, Mary W. Galinski,

; APPLICANT: Samuel P. Wertheimer

; TITLE OF INVENTION: MEROZOITE ANTIGENS LOCALIZED AT

; TITLE OF INVENTION: THE APICAL END OF THE PARASITE

; NUMBER OF SEQUENCES: 3

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Darby & Darby P.C.

; STREET: 805 Third Avenue

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10022

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 5.25 inch, 360 Kb storage

; COMPUTER: IBM or IBM-compatible

; OPERATING SYSTEM: PC/MS-DOS

; SOFTWARE: Wordperfect

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/792,865D

; FILING DATE: 19911004

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/US90/01849

; FILING DATE: April 3, 1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Jackson, Arthur

; REGISTRATION NUMBER: 34,354

; REFERENCE/DOCKET NUMBER: 5986/14692-US2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 527-7700

; TELEFAX: (212) 753-6237

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3763 base pairs


```
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; DESCRIPTION:
; HYPOTHETICAL: No
; ANTI-SENSE: No
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: P.vivax
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE:
; ORGANELLE:
; IMMEDIATE SOURCE:
; LIBRARY: Lambda gt 11 native P.vivax
; LIBRARY: DNA expression library
; CLONE: 5.3
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: This sequence corresponds to
; OTHER INFORMATION: Figure 1A (sheets 1-4) in the
; OTHER INFORMATION: application, as filed.
; PUBLICATION INFORMATION:
; AUTHORS:
; TITLE:
; JOURNAL:
; VOLUME:
; ISSUE:
; PAGES:
; DATE:
; DOCUMENT NUMBER:
; FILING DATE:
; PUBLICATION DATE:
; RELEVANT RESIDUES IN SEQ ID NO:
; US-07-792-865D-1

Alignment Scores:
Pred. No.: 245 Length: 3763
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 1 Gaps: 0

US-10-751-736-84 (1-907) x US-07-792-865D-1 (1-3763)

QY 730 IleLeuLeuAsnSerLeuCysPheLeu 738
Db 2576 ATTTGCTTAACTCACTATGTTCTTG 2550

Search completed: July 12, 2005, 13:02:23
Job time : 436 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 12, 2005, 07:34:33 ; Search time 33 Seconds
(without alignments)
2051.718 Million cell updates/sec

Title: US-10-751-736-84
Perfect score: 4702
Sequence: 1 MDTSLGVLVLLPVLQLAT.....PAYPTESCHLSSVAFVCL 907

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/protdata/1/1aa/5A-COMB.pep:*
2: /cgn2_6/protdata/1/1aa/5B-COMB.pep:*
3: /cgn2_6/protdata/1/1aa/6A-COMB.pep:*
4: /cgn2_6/protdata/1/1aa/6B-COMB.pep:*
5: /cgn2_6/protdata/1/1aa/PCTUS-COMB.pep:*
6: /cgn2_6/protdata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4702	100.0	907	4	US-09-170-496D-264
2	4596	99.9	907	4	US-09-170-496D-278
3	1217.5	25.9	644	2	US-08-866-757-2
4	1217.5	25.9	644	3	US-09-153-593-2
5	636	13.5	692	3	US-07-757-342D-6
6	636	13.5	692	4	US-09-461-657B-6
7	621.5	13.2	695	1	US-08-487-886-2
8	621.5	13.2	695	3	US-08-482-855-2
9	621.5	13.2	695	3	US-08-474-986-2
10	598	12.7	764	3	US-07-741-453A-61
11	595	12.7	764	3	US-07-757-342D-5
12	595	12.7	764	4	US-09-461-657B-5
13	595	12.7	1181	4	US-09-826-509-587
14	594	12.6	764	3	US-07-741-453A-59
15	593	12.6	764	4	US-09-826-509-403
16	591	12.6	764	3	US-07-741-453A-54
17	588.5	12.5	700	3	US-07-757-342D-3
18	588.5	12.5	700	4	US-09-461-657B-3
19	588	12.5	764	4	US-09-826-509-399
20	588	12.5	764	4	US-09-826-509-407
21	586	12.5	764	4	US-09-826-509-395
22	586	12.5	764	4	US-09-826-509-415
23	579	12.3	764	4	US-09-826-509-411
24	575	12.2	764	4	US-09-826-509-419
25	574	12.2	696	3	US-07-757-342D-4
26	574	12.2	696	4	US-09-461-657B-4
27	573	12.2	764	4	US-09-826-509-423

28	569	12.1	764	3	US-07-741-453A-60	Sequence 60, Appl
29	565	12.0	764	4	US-09-186-350A-53	Sequence 53, Appl
30	560	11.9	699	3	US-07-757-342D-2	Sequence 2, Appl
31	560	11.9	699	4	US-09-461-657B-2	Sequence 2, Appl
32	550.5	11.7	674	3	US-07-757-342D-10	Sequence 10, Appl
33	550.5	11.7	674	4	US-09-461-657B-10	Sequence 10, Appl
34	503.5	10.7	795	3	US-07-741-453A-55	Sequence 55, Appl
35	499.5	10.6	792	3	US-07-741-453A-56	Sequence 56, Appl
36	471	10.0	764	3	US-07-741-453A-29	Sequence 29, Appl
37	469.5	10.0	636	3	US-07-757-342D-7	Sequence 7, Appl
38	469.5	10.0	636	4	US-09-461-657B-7	Sequence 7, Appl
39	460	9.8	611	3	US-07-757-342D-8	Sequence 8, Appl
40	460	9.8	611	4	US-09-461-657B-8	Sequence 8, Appl
41	426.5	9.1	603	1	US-08-190-802A-50	Sequence 50, Appl
42	426.5	9.1	603	3	US-08-477-346-50	Sequence 50, Appl
43	426.5	9.1	603	4	US-08-473-089-50	Sequence 50, Appl
44	426.5	9.1	603	4	US-08-487-072A-50	Sequence 50, Appl
45	421	9.0	605	3	US-09-063-950-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-09-170-496D-264
; Sequence 264, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 264
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-264

Query Match	100.0%	Score 4702;	DB 4;	Length 907;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 907;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDTSLGVLVLLPVLQLATGSGSPRSGLVLRGCPHCHCEPDGRMLLRVDCSDGLGSEL	60	
Db	1	MDTSLGVLVLLPVLQLATGSGSPRSGLVLRGCPHCHCEPDGRMLLRVDCSDGLGSEL	60	
QY	61	PSNLSVFTSYLDLSNNNISQLPNPLPSLRFLFEEURLAGNALTYIPKGAFTGLYSKVLUM	120	
Db	61	PSNLSVFTSYLDLSNNNISQLPNPLPSLRFLFEEURLAGNALTYIPKGAFTGLYSKVLUM	120	
QY	121	LQNNQLRHPVPEALQNLRLDANHSIVPPSCFSGLSRLHRLMDNALTEIPVQ	180	
Db	121	LQNNQLRHPVPEALQNLRLDANHSIVPPSCFSGLSRLHRLMDNALTEIPVQ	180	
QY	181	AFRSLSALQAMTALANKITHIIPDYAFGNLSLVLHNNRHSIGKCKFCFGLHSLETID	240	
Db	181	AFRSLSALQAMTALANKITHIIPDYAFGNLSLVLHNNRHSIGKCKFCFGLHSLETID	240	
QY	241	LYNNNLDEFFTAIRTLNLKELGPHSNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA	300	
Db	241	LYNNNLDEFFTAIRTLNLKELGPHSNNIRSIPEKAFVGNPSLTIHFYDNP IQVGRSA	300	
QY	301	FOHLPELRTLNGASQITTEPDLTGATNLESLLTGQOISLPTVCNQLPNLOVLDS	360	
Db	301	FOHLPELRTLNGASQITTEPDLTGATNLESLLTGQOISLPTVCNQLPNLOVLDS	360	
QY	361	YNLLEDLPSFVSCQKQKIDLRHNEIYEIKVDTFQQLSLRSLNLAWNKIAIHPNAPST	420	

361	Db	YNLLEDLPFSVCQKLOKIDLRHNEIYEIKVDYTFQQLLSRLSNLAWNKTAIITHPAFST	420
421	Qy	LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLJSSNFPELKVIEMPYAYQCC	480
421	Db	LPSLIKLDLSSNLLSFPITGLHGLTHLKTGNHALQSLJSSNFPELKVIEMPYAYQCC	480
481	Qy	AFGVCEPAYKISNQNWKGDNSMDDLHKDKDAGMFOQOERDLEDFLDPEEDLKALHSVQ	540
481	Db	AFGVCEPAYKISNQNWKGDNSMDDLHKDKDAGMFOQOERDLEDFLDPEEDLKALHSVQ	540
541	Qy	CSPSPGPFKPCBHLLDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYTSPIKLLIGVIA	600
541	Db	CSPSPGPFKPCBHLLDGMLIRIGVWTIAVLATCNALVTSTVFRSPLYTSPIKLLIGVIA	600
601	Qy	AVNMLTGVSSAVLAGVDRAFTGSPFARHGAWWENGVCCHVIGTFLSIFASESSVFLLTAAAL	660
601	Db	AVNMLTGVSSAVLAGVDRAFTGSPFARHGAWWENGVCCHVIGTFLSIFASESSVFLLTAAAL	660
661	Qy	ERGFSVKYSAKFETKAPSSSLKVIILLCALLATMAAVPILGGSKYCASPLCLPLPFGBP	720
661	Db	ERGFSVKYSAKFETKAPSSSLKVIILLCALLATMAAVPILGGSKYCASPLCLPLPFGBP	720
721	Qy	STWGYMVALILNSLCFLMMTITAYTKLCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780
721	Db	STWGYMVALILNSLCFLMMTITAYTKLCNLDKGDLENIWDCSMVKHIALLLFTNCILNC	780
781	Qy	PVAFISFSSLINLTFTISPEVIFKILLVVVPLPACINPLLTLFNPFPKEDLVSLRKQTVV	840
781	Db	PVAFISFSSLINLTFTISPEVIFKILLVVVPLPACINPLLTLFNPFPKEDLVSLRKQTVV	840
841	Qy	WTRSKHPSLMSINSDDEKQSCDSQALVTFTSSITVDLPSSVPSPAYPVVTTESCHLS	900
841	Db	WTRSKHPSLMSINSDDEKQSCDSQALVTFTSSITVDLPSSVPSPAYPVVTTESCHLS	900
901	Qy	VAFPVCL 907	
901	Db	VAFPVCL 907	

RESULT 2

```

US-09-170-496D-278      ; Sequence 278, Application US/09170496D
; Patent No. 655339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 655339-Endog
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 278
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-170-496D-278

```

Query Match 99.9%; Score 4696; DB 4; Length 907;

Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 906:	Conservative	0;	Mismatches 1;
Indels	0;	Gaps	0;

[illegible]

Qy	121	LQNNQLRHHVPTREALQNLRSLOSLRLDANHISYVPPSCFSGHLSLRLDNDNALTEIPVQ	180
Db	121	LQNNQLRHHVPTREALQNLRSLOSLRLDANHISYVPPSCFSGHLSLRLDNDNALTEIPVQ	180
Qy	181	AFRSLSALOQWTLALNKIHHIPDYAFGNLSLVLVHLHNNRIHSGKKCFGLHSELETLD	240
Db	181	AFRSLSALOQWTLALNKIHHIPDYAFGNLSLVLVHLHNNRIHSGKKCFGLHSELETLD	240
Qy	241	LYNNNLDEFPPTAIRTLNLSKELGPHSNNIRSIPEKAFVGNPFSLTIHFYDNPICPVGRSA	300
Db	241	LYNNNLDEFPPTAIRTLNLSKELGPHSNNIRSIPEKAFVGNPFSLTIHFYDNPICPVGRSA	300
Qy	301	FOHLPRLRTLTLNGASQITTEPDDLTGTANLESRLTGAQISSLPQTVCNQPLNLOVLDSL	360
Db	301	FOHLPRLRTLTLNGASQITTEPDDLTGTANLESRLTGAQISSLPQTVCNQPLNLOVLDSL	360
Qy	361	YNLLEDLPSSVCOKLOKIDLRHNEIYEIKVDTTQQLLSLPSLANLAWNKIAIHHNPAEST	420
Db	361	YNLLEDLPSSVCOKLOKIDLRHNEIYEIKVDTTQQLLSLPSLANLAWNKIAIHHNPAEST	420
Qy	421	LPSLIKLDLSNLLSSPPIITGLHGLTHLKTGNHALQSLISSENPPELKVIEMPAYQCC	480
Db	421	LPSLIKLDLSNLLSSPPIITGLHGLTHLKTGNHALQSLISSENPPELKVIEMPAYQCC	480
Qy	481	AFGCENAYKISNOWNKGDSSMDDLHKKDAGMFOAQDERDLEPLLDPEBDLKALHSVQ	540
Db	481	AFGCENAYKISNOWNKGDSSMDDLHKKDAGMFOAQDERDLEPLLDPEBDLKALHSVQ	540
Qy	541	CSPSPGPFKPECHLLDQWLIBRIGVWTTAVLATCNALVTSTVFRSPLYISIPKLLIGVIA	600
Db	541	CSPSPGPFKPECHLLDQWLIBRIGVWTTAVLATCNALVTSTVFRSPLYISIPKLLIGVIA	600
Qy	601	AVNNMLTGVSAAVLAGVDAFTGSPARHGAWENGVCVHVGFLSIFASESSVFLTLAAL	660
Db	601	AVNNMLTGVSAAVLAGVDAFTGSPARHGAWENGVCVHVGFLSIFASESSVFLTLAAL	660
Qy	661	ERGSFVKYSAKFETKAPFSSLKVIILLCALLALMAAAVPLLGGSKYGASPLCLPFGPEP	720
Db	661	ERGSFVKYSAKFETKAPFSSLKVIILLCALLALMAAAVPLLGGSKYGASPLCLPFGPEP	720
Qy	721	STMGYMWALILLNSLCFLMWTIATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC	780
Db	721	STMGYMWALILLNSLCFLMWTIATKLYCNLDKGDLENWDCSMVKHIALLLFTNCILNC	780
Qy	781	PVAFLPSSLLNLNFTFISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTVV	840
Db	781	PVAFLPSSLLNLNFTFISPEVIKFIILLVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTVV	840
Qy	841	WTRSKHPFLMSINSDDVEKQSCDSTQALVTFTSSSITYDLPPSSVPSPAYPVTESCHLSS	900
Db	841	WTRSKHPFLMSINSDDVEKQSCDSTQALVTFTSSSITYDLPPSSVPSPAYPVTESCHLSS	900
Qy	901	VAFVPCL 907	
Db	901	VAFVPCL 907	
RESULT 3			
US-08-866-757-2			
; Sequence 2, Application US/08966757			
; Patent No. 5858716			
; GENERAL INFORMATION:			
; APPLICANT: ELSHOURBAGY, NABIL A			
; APPLICANT: LI, XIAOTONG			
; APPLICANT: BERGMSA, DERK J			
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)			
; NUMBER OF SEQUENCES: 4			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: RATNER & PRESTIA			
; STREET: P.O. BOX 980			
; CITY: VALLEY FORGE			
; STATE: PA			
; COUNTRY: USA			


```

; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 644 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-866-757-2

```

```

Query Match      25.9%; Score 1217.5; DB 2; Length 644;
Best Local Similarity 46.3%; Pred. No. 6.1e-89;
Matches 250; Conservative 88; Mismatches 183; Indels 19; Gaps 5;

QY 318 ITPEDLTGTANLESILTTGAQISSLPQVCNQLPNQLVDLSYNNLEDLPFSVCQKIQ 377
Db 2 VQOFPNLGTGVHLESILTTGTAKISSIPNNLCQEQKMLRTLDLSYNNIRDLPSFNGCHALE 61

QY 378 KIDLRHNEIYEIKVDYTFQQLSLRLSNLAWNKIAIHPNAPSTLPKLDLSSNLLSSF 437
Db 62 EISLQRNQIYQKEGTFQGLISRLDLRNLHIHSRAFATLGPITNLDVSPNELTSF 121

QY 438 PITGLHGLTHLKTGNHALQSLISSENPPELKIEMPYAYQCCAFGVCENAYKISQNWKN 497
Db 122 PTEGLNGLNQLKLVGNFKLKEALAAKDFVNLRLSLVPYAYQCCAFWGCDSYANL---NT 177

QY 498 GDNSSMDLHKDAGMFOQDERDLEDFLDPEEDLKALHS---VQCSPPGPKPCHEHL 554
Db 178 EDNSLQDHSVQAEKGTADAANVT-----TLENEE-----HSQIIHCTPSTGAFKPCBYL 228

QY 555 LDGWLIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIAAVNMLTGVSASVLA 614
Db 229 LGSWIRLTVWFIFLVALFFNLLVLTTFASCTSLPSSKLFGLISVNLPMGIYTGILT 288

QY 615 GVDATFGSFARHGAWENGVCVIGFSLIPASESSVFLTLAALERGFVSYSKAPET 674
Db 289 FLDAVSWGRFAEFGLWWTGSGCKVTGFLAVPSSSAIFLLMLATVERSLAKOIMKNGK 348

QY 675 KAPFSSLVIIILLCALLTMAAVPLLGSKYKASPLCLPLPFGEPTMTGYMVALILNS 734
Db 349 SNHLKQFRVAALLAFLGATVTCGPFPHRGYSASPLCLPPTGTGETPSLGTFTVLVLNS 408

QY 735 LCFLLMTIATYKLYCNLDKGLDENIWDGSMVKHIALFTNCILNCIPVAFLSFSSLINLT 794
Db 409 LAFLLMAYITYKLYCNLEKEDLSNSQSMIKHVAWLIFTNCIFPCPVAFPSFAPLITAI 468

QY 795 FTSPEVIKFIILVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYVYVTRSKHPSLMSINS 854
Db 469 SISPEIMKSVTLIFFPLPACLNPLVLYFNFNPKFKEDWKLKRRV---TKKSGSVSVSISS 525

```

RESULT 4

US-09-153-593-2

```

; Sequence 2, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 644
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-153-593-2

```

```

Query Match      25.9%; Score 1217.5; DB 3; Length 644;
Best Local Similarity 46.3%; Pred. No. 6.1e-89;
Matches 250; Conservative 88; Mismatches 183; Indels 19; Gaps 5;

QY 318 ITPEDLTGTANLESILTTGAQISSLPQVCNQLPNQLVDLSYNNLEDLPFSVCQKIQ 377
Db 2 VQOFPNLGTGVHLESILTTGTAKISSIPNNLCQEQKMLRTLDLSYNNIRDLPSFNGCHALE 61

QY 378 KIDLRHNEIYEIKVDYTFQQLSLRLSNLAWNKIAIHPNAPSTLPKLDLSSNLLSSF 437
Db 62 EISLQRNQIYQKEGTFQGLISRLDLRNLHIHSRAFATLGPITNLDVSPNELTSF 121

QY 438 PITGLHGLTHLKTGNHALQSLISSENPPELKIEMPYAYQCCAFGVCENAYKISQNWKN 497
Db 122 PTEGLNGLNQLKLVGNFKLKEALAAKDFVNLRLSLVPYAYQCCAFWGCDSYANL---NT 177

QY 498 GDNSSMDLHKDAGMFOQDERDLEDFLDPEEDLKALHS---VQCSPPGPKPCHEHL 554
Db 178 EDNSLQDHSVQAEKGTADAANVT-----TLENEE-----HSQIIHCTPSTGAFKPCBYL 228

QY 555 LDGWLIRIGVWTTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIAAVNMLTGVSASVLA 614
Db 229 LGSWIRLTVWFIFLVALFFNLLVLTTFASCTSLPSSKLFGLISVNLPMGIYTGILT 288

QY 615 GVDATFGSFARHGAWENGVCVIGFSLIPASESSVFLTLAALERGFVSYSKAPET 674
Db 289 FLDAVSWGRFAEFGLWWTGSGCKVTGFLAVPSSSAIFLLMLATVERSLAKOIMKNGK 348

QY 675 KAPFSSLVIIILLCALLTMAAVPLLGSKYKASPLCLPLPFGEPTMTGYMVALILNS 734
Db 349 SNHLKQFRVAALLAFLGATVTCGPFPHRGYSASPLCLPPTGTGETPSLGTFTVLVLNS 408

QY 735 LCFLLMTIATYKLYCNLDKGLDENIWDGSMVKHIALFTNCILNCIPVAFLSFSSLINLT 794
Db 409 LAFLLMAYITYKLYCNLEKEDLSNSQSMIKHVAWLIFTNCIFPCPVAFPSFAPLITAI 468

QY 795 FTSPEVIKFIILVVVPLPACLNPLLYILFNPHFKEDLVSLRKQTYVYVTRSKHPSLMSINS 854
Db 469 SISPEIMKSVTLIFFPLPACLNPLVLYFNFNPKFKEDWKLKRRV---TKKSGSVSVSISS 525

```

RESULT 5

US-07-757-342D-6

```

; Sequence 6, Application US/07757342D
; Patent No. 6218509
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, Masao
; APPLICANT: MINEGISHI, Takashi
; APPLICANT: NAKAMURA, Kazuto
; TITLE OF INVENTION: PROTEIN, DNA AND USE THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &

```



```
; ; CUSHMAN
; STREET: 130 Water Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: US
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/757,342D
; FILING DATE: 10-Sep-1991
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BUCKLEY, Linda M.
; REGISTRATION NUMBER: 31003
; REFERENCE/DOCKET NUMBER: 41226
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)523-3400
; TELEFAX: (617)523-6440
; TELEX: 200291 STRE UR
;
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 692 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
; MOLECULE TYPE: protein
;
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-07-757-342D-6

Query Match 13.5%; Score 636; DB 3; Length 692;
Best Local Similarity 23.1%; Pred. No. 3.1e-42;
Matches 211; Conservative 129; Mismatches 310; Indels 264; Gaps 25;

Qy 9 LLSPLVLIQLATGSSPRSGVLLRGCPH---CHCEPDGMLLRVDCSDGLSELPSNLS 65
Db 3 LLLVSLLAFLGTGS-----GC--HHWLCHC--SNRVFL---CQDSKVTEIPT--- 42

Qy 66 VFTSYLDLSMNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFGLYSL-KVLMQLNN 124
Db 43 -----DLPRNAI-----ELRFVLTCLRVPKGSFAGDLEKIEISQND 81

Qy 125 QLRHVPTALQNLRSQSLRD-ANHISYVPPSCFSGHLSLRHLWLDNALTEIPVQAFR 183
Db 82 VLEVEADVFNLPKLHEIRIEKANNLLYINPEAFONPLPSRLYLLISNTGKHLF--AVH 139

Qy 184 SLSALQMTALANKIHHIPDYAFGNLSLVLHLHNNRIHSLGKCKFDGLHSLETLDLNY 243
Db 140 KIQSLQKVLDDIQ-----DNINIHIVARNSMFGL-----168

Qy 244 NNLDFFPTAIRTLNKLGHFSNNIRSIPEKAFVGNPSLTIHFYDNPIQVGRSAFOH 303
Db 169 ----SFESVILWLS-----KNGIEIHNCFAFG-----192

Qy 304 LPRLRTLNGASQITTEPPDLTGNTANLESFLTGTGAQISSLPQTVCNQLPNLOVLDSLNL 363
Db 193 -----TQDELNLSD--NNN 205

Qy 364 LEDLPFSVCQKQKIDLRHNEIYEIKVDFTFQQLLSRLSLNLANWKIAIHPNAPSTLPS 423
Db 206 LEELPN-----DVFOGASG 219

Qy 424 LTKLDSNLLSSPFIITGLHGLTHLKLGNHALQSLISSENPPELKVTEMPIYQCCAFG 483
Db 220 PVILDIRTKVHSLPNHGLENLKCLRARSYRLKXLPNLDKFTVLMEASLTVPSHCCAF 279

Qy 484 -----VCENA-----YKISQWNGKDN--SSMDP--LHKDAGNFQOQDERLEDLF 525
Db 280 NLKROISELPHCNKSIIRQDIDDMTQIGDQVSLIDDPSSYKSGSDMMYNEFDYDLGNE 339

Qy 526 LLDPEEDLKALHSVCSPSPGPKCEHLLDGLWLRIGVWTIATVALATCNALVTSTVFRS 585
```

```
Db 340 VWD-----VTCSPKPDFAFNPCEDIMGYNILRVLIWFISILAITGNTTVLVLTTS 389
Qy 586 PLYISPIKLLIGVITAAVNMLTGVSSAVLAGVDAPTFGSPARHGAWWENGCVHIGFLSI 645
Db 390 QYKLTVPFRFLMCLNLAFLADLCIGIYLLLIASVDIHTKSQYHNYAIDWQTAGCDAAGFTV 449
Qy 646 FASESSVFLTLAALERGFVSVKYSAKETKAPSSSLKVIILLCALLTAAWAVPLLGSK 705
Db 450 FASELSVYTLTATLERWHTITHAMQLECKVQURHAASVMVLTGWTFAAALPFIPIGLSS 509
Qy 706 YGASPLCLPLPFGPESTMGYMVALILNLSLCFLAMTIATYTKLYCNLDKGD-ENIMDCSM 764
Db 510 YMKVSIICLPMDIDSPLSQLYVMALLVNLVAFVVICCYTHIYLTVRNPITVSSSDTKI 569
Qy 765 VKHIALLLFTNCILNCPVAFSLFSSSLNLTIFISPEVIKFIILLVVVPLPACLNPLLYLFN 824
Db 570 AKRMATLIFTDFLCMAPISPFAISASLKVPILITVSKAKILLVLFYFIPINSCANPFLYAIT 629
Qy 825 PHFKEDLVSL-----RKQTYWTRSKHPSLMSINSDDVEKQSCDSTQALVT 870
Db 630 KNFRDFFILLSKPGCYEMQAQIYRTTSSATHNFH-----ARKSHCSAPRV-- 677
Qy 871 FTSSSITYDLPPSS 884
Db 678 -TNSYVLVPLNHSS 690

RESULT 6
US-09-461-657B-6
; Sequence 6, Application US/09461657B
; Patent No. 6635445
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, MASAO
; APPLICANT: MINEGISHI, TAKASHI
; APPLICANT: NAKAMURA, KAZUTO
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN LUTEINIZING
; TITLE OF INVENTION: HORMONE-HUMAN CHORIONIC GONADOTROPIN RECEPTOR PROTEIN
; TITLE OF INVENTION: AND TRANSFORMANTS THEREOF
; FILE REFERENCE: 70281/41,226C
; CURRENT APPLICATION NUMBER: US/09/461,657B
; CURRENT FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: 07/757,342
; PRIOR FILING DATE: 1991-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 6
; LENGTH: 692
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-461-657B-6

Query Match 13.5%; Score 636; DB 4; Length 692;
Best Local Similarity 23.1%; Pred. No. 3.1e-42;
Matches 211; Conservative 129; Mismatches 310; Indels 264; Gaps 25;

Qy 9 LLSPLVLIQLATGSSPRSGVLLRGCPH---CHCEPDGMLLRVDCSDGLSELPSNLS 65
Db 3 LLLVSLLAFLGTGS-----GC--HHWLCHC--SNRVFL---CQDSKVTEIPT--- 42

Qy 66 VFTSYLDLSMNNISQLLPNPLPSLRFLEELRAGNALTYIPKGAFGLYSL-KVLMQLNN 124
Db 43 -----DLPRNAI-----ELRFVLTCLRVPKGSFAGDLEKIEISQND 81

Qy 125 QLRHVPTALQNLRSQSLRD-ANHISYVPPSCFSGHLSLRHLWLDNALTEIPVQAFR 183
Db 82 VLEVEADVFNLPKLHEIRIEKANNLLYINPEAFONPLPSRLYLLISNTGKHLF--AVH 139

Qy 184 SLSALQMTALANKIHHIPDYAFGNLSLVLHLHNNRIHSLGKCKFDGLHSLETLDLNY 243
Db 140 KIQSLQKVLDDIQ-----DNINIHIVARNSMFGL-----168

Qy 244 NNLDFFPTAIRTLNKLGHFSNNIRSIPEKAFVGNPSLTIHFYDNPIQVGRSAFOH 303
```



```
Db 169 -----SFESVILWLS-----KNGIEETHNCAFNG-----192
Qy 304 LPELRTLTLNGASQITEPDLTGATNLESITLTGAQISLSPOTVCNQLPNQVLDLSVNL 363
Db 193 -----TQDELNLSD--NNN 205
Qy 364 LBDLPSPVQCKLOKIDLRHNEIYKVDTFQOOLSLRSLNLANWKIAIHPNAPSTLPS 423
Db 206 LBEELN-----DVEFGASG 219
Qy 424 LKLDLSSNLLSPPTGLHGLTHLKTGNHALQSLISSENPPELKIEMPYAYOCCARG 483
Db 220 PVLDISRTKXSLNHLNGLNKLKLRARSTYRKLPLNDKFTLMEASLTYPSPHCCAPA 279
Qy 484 -----VCENA---YKISNOWNKGDN--SSMDD--LHKDQAGMFOAQDERDLEDF 525
Db 280 NLKQRISEHLPCNKISILLRQDIDDMTQIGDQVSLIDDEPSYKSGSDMMYNEFDYDLNE 339
Qy 526 LLDPFEDLKALHSVQSPSPGPKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRS 585
Db 340 VVD-----VTCSPKPDAPNCPEDIMGYNILRLVLIWFISILAITGNTTTLVVLVLTTS 389
Qy 586 PLYISPIKLLIGVIAVNMVLTGVSSNAVLAGVDAPFGSPARHGANWENGCHVIGPLSI 645
Db 390 QYKLTVPRELMCNLAPADLCIGIYLLLIASVDIHTKSQYHNYAIDWQTGAGCDAAGFTV 449
Qy 646 FASESSVLLTLAALRERGSVKYSAKFETKAPFSSLKVILLCALLTMAAIVPLLGSK 705
Db 450 FASELSVYTLTALTERTWHITTHAQLECKVQRHAAVWLVGTFAPAAALFPFIFGLSS 509
Qy 706 YGASPLCLPLPFGPESTMGVVALILNLSCLFMMTIAYTKLYCNLDKGLD-ENIWDCSM 764
Db 510 YMKVSICLPMDIDSPLSQLYVMAVLLVNLVAFVWICGVTHIYLTVRNPTIVSSSDTKI 569
Qy 765 VGHIALLLTNCILNCPVAFSSSLINLTFISPEVIKILLVVPPLACINPLIYLFPN 824
Db 570 AKRMATLIPTDFLCMAPISFPAISASLKVPLITVSKAKILLVLFYPIINSCANPFLYAIPT 629
Qy 825 PHFKEDLVSL-----RKOTVWTRSKHPSLMSINSDVKEQSCDSTQALVT 870
Db 630 KNFRDRFFLLSKFGYENQAOIYRTETSSATNTH-----ARKSHCSSAPRV-- 677
Qy 871 FTSSITYDLPPSS 884
Db 678 -TNSYVLVPLNHSS 690

RESULT 7
US-08-487-886-2
; Sequence 2, Application US/08487886
; Patent No. 574448
; GENERAL INFORMATION:
; APPLICANT: Kelton, Christie Ann
; APPLICANT: Schweickhardt, Rene Lynn
; APPLICANT: Cheng, Shirley Vui Yen
; APPLICANT: Nugent, No. 574448een Patrice
; TITLE OF INVENTION: Human Follicle Stimulating
; TITLE OF INVENTION: Hormone Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephan P. Williams,
; ADDRESSEE: Ares-Serono, Inc.
; STREET: Exchange Place, 37th floor
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
; COMPUTER: IBM PS/2, model 55 SX
; OPERATING SYSTEM: MS-DOS version 4.0
; SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
```

```
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,886
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/670,085
; FILING DATE: 15-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Stephan P.
; REGISTRATION NUMBER: 28546
; REFERENCE/DOCKET NUMBER: US/252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 723-1300
; TELEFAX: (617) 723-8923
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 695
; TYPE: Amino acid
; TOPOLOGY: Linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: signal sequence
; LOCATION: -17 to -1
; IDENTIFICATION METHOD: hydrophobic
; FEATURE:
; NAME/KEY: putative amino-terminal extracellular domain
; LOCATION: 1 to 349
; IDENTIFICATION METHOD: similarity with other
; IDENTIFICATION METHOD: dimeric glycoprotein receptor extracellular
; IDENTIFICATION METHOD: domains, hydrophilic
; FEATURE:
; NAME/KEY: transmembrane domain
; LOCATION: 350 to 613
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane domains
; FEATURE:
; NAME/KEY: putative transmembrane region I
; LOCATION: 350 to 370
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region II
; LOCATION: 382 to 404
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region III
; LOCATION: 427 to 448
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region IV
; LOCATION: 469 to 491
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region V
; LOCATION: 512 to 533
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region VI
; LOCATION: 557 to 580
; IDENTIFICATION METHOD: similarity to other G
; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
; FEATURE:
; NAME/KEY: putative transmembrane region VII
```



```
;; NAME/KEY: putative transmembrane region II
;; LOCATION: 382 to 404
;; IDENTIFICATION METHOD: similarity to other G
;; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
;; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
;; FEATURE:
;; NAME/KEY: putative transmembrane region III
;; LOCATION: 427 to 448
;; IDENTIFICATION METHOD: similarity to other G
;; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
;; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
;; FEATURE:
;; NAME/KEY: putative transmembrane region IV
;; LOCATION: 469 to 491
;; IDENTIFICATION METHOD: similarity to other G
;; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
;; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
;; FEATURE:
;; NAME/KEY: putative transmembrane region V
;; LOCATION: 512 to 533
;; IDENTIFICATION METHOD: similarity to other G
;; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
;; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
;; FEATURE:
;; NAME/KEY: putative transmembrane region VI
;; LOCATION: 557 to 580
;; IDENTIFICATION METHOD: similarity to other G
;; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
;; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
;; FEATURE:
;; NAME/KEY: putative transmembrane region VII
;; LOCATION: 592 to 613
;; IDENTIFICATION METHOD: similarity to other G
;; IDENTIFICATION METHOD: protein-coupled receptor transmembrane regions,
;; IDENTIFICATION METHOD: hydrophobic, about 20-23 amino acids in length
;; FEATURE:
;; NAME/KEY: putative carboxy-terminal intracellular
;; NAME/KEY: domain
;; LOCATION: 614 to 678
US-08-482-855-2

Query Match 13.2%; Score 621.5; DB 3; Length 695;
Best Local Similarity 22.4%; Pred. No. 4.5e-41;
Matches 207; Conservative 146; Mismatches 306; Indels 265; Gaps 25;

QY 8 VLLSLPVLQLATGSSPRSVLLRGCPHCH---CEPDGRMLLRVDCSDLGSLSPNL 64
Db 4 LLVSLAFSLGSG-----CHRIHCNSRVFL---COESKVTIPS-- 42
QY 65 SVFTSYLDLSMNNISQLLPNPLSLRFLLEELRAGNALTYPKGAPTLGLYSL-KVLMLO 123
Db 43 -----DLPRNAI-----ELRFVLTCLRVIQKGAFGDLEKTEISQN 80
QY 124 NQLRHVPTEANQLRSQSLRLD-ANHISYVPPSCFSGHLSHLHLLDDNALTEIPVOAF 182
Db 81 DVLVEIADVFNLPKJHEIRIEKANLLYINP-----BAP 116
QY 183 RSLSLAQMTALNKIHHIPDVFAGNLSLVVHLHNNRIHSLGKKCFDGLHSLRLELDLN 242
Db 117 QNLPNQLVLLISNTGKHLPDV-----HKHSLOK-----VLLDIQ 152
QY 243 YNNLDEFPPTAINTLSNLKELGPHSNIRSIPEKAFVGNPSLTIHFYDNPQFVGRSAFQ 302
Db 153 DN-----ININTIERNSPVGLS-----FESVILWLNKNGIQ 183
QY 303 HLPRLTILNCASQITEFPDLTGANLESILTCAQISSLPQTCNQLPNLOVLDLSYN 362
Db 184 ---ETHCAFNG-----TQDELNLSD--NN 204
QY 363 LLEDLPSPSVQCQLQKIDLRHNEIYEIKVDITFQQLLSRLSLNANWKIAIHPNAPSTLP 422
Db 205 NLEELPN-----DVFHGAS 218
```

```
QY 423 SLIKDLSSNLLSPPTIGLHGLTHLKLTHGNHALQSLLISSENFPELKVIEPMPAYQCCAP 482
Db 219 GFVILDISTRIRHSLSYGLNLLKGLRARSTYNLXKGLPTLEKLVAMLEASLYPSHCCAF 278
QY 483 G-----VCBNAYKISNQ-----MNGKDNSSMDLDLHKDAGMFOAQDERDLEDF 525
Db 279 ANWRRIQISELHPICNKS--ILRQVDVYMTQTRGQRSSLAEDN-----ESSYSRGFDMT 329
QY 526 LLDFEEDL-KALHSVQCSPPSPGPKPCBHLDDGMLIRIGVMTIAVLATCNALVTSTVPR 584
Db 330 YTEFDYDLCEVNDVTCSPKDAFNPCEIDIMGYNLRVLINWFISILAITGNIIVLILTT 389
QY 585 SPLYSPIKLILIGVIAAVNMLTGVSSAVLAGVDAFTFGSPARHGAWWENGCVCHVIGLS 644
Db 390 SOYKLTVPRLWMLNLAFAADLCIGIYLLIASVDIHTKSQYHNYAIDWQTGAGCDAAEFT 449
QY 645 IPASESSVFLATLAALERGFVSYSKAPETKAPSSSLKVIILLCALLATMAAVPLLGGS 704
Db 450 VFASLSVYTLTATLERWHTITHAMQLDCKVQLRHAASVMVMGWIFAPAAALFFIGIS 509
QY 705 KYGASPLCLPFPGEPSPTMGVMYAILILNSICFLMMTIATYKLYCNLDKGLD-ENIWDCS 763
Db 510 SYMKVSIICLPMDIDSPLSQLYVMSLLVNLVAFVVICGYTHIYLTVRNPNIIVSSSDTR 569
QY 764 MVKHIALLLFTNCILNCPVAFSLSPSSLINLTFISPEVIKFIILLVVVPLPACINPLLYLF 823
Db 570 IAKRMALIFTDFLCMAPISFPAISASLUKPLITVSKAKILLVLPHPINSCANPELYAIF 629
QY 824 NPHFKEDLVLSRKQTYVWTRSKHPSLMSINGDDVEKQSCDSTQALV--TFTSSSITYDLP 881
Db 630 TKNFRDFFIL-----LSKCCYEMQAQIYETEISSTVHTHP 667
QY 882 PSSVPSPAYPVTESCHLSSVAFVP 905
Db 668 RNHCSSAPRVNG---STYLVP 688

RESULT 9
US-08-474-986-2
; GENERAL INFORMATION:
; APPLICANT: Kelton, Christie Ann
; Schweickhardt, Rene Lynn
; Cheng, Shirley Vui Yen
; Nugent, No. 63721leen Patrice
; TITLE OF INVENTION: Human Folicle Stimulating
; Hormone Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Stephen P. Williams,
; Ares-Serono, Inc.
; STREET: Exchange Place, 37th floor
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" diskette, 1.44 MB, high density
; COMPUTER: IBM PS/2, model 55 SX
; OPERATING SYSTEM: MS-DOS version 4.0
; SOFTWARE: VAX/VMS Mass11 via Kermit to IBM MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474,986
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/670,085
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams, Stephan P.
; REGISTRATION NUMBER: 28546
; REFERENCE/DOCKET NUMBER: US/252
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 723-1300
```


Qy 720 PSTGMVVALILLNSLCLFMMTIATYKLYCNL-----DKGLENIDWCSMVKHIALLLFT 774
Db 577 PLALAYIVFVLINIAVAVIVCCCVKIYITVRNPQYNGDK-----DTKIARMAVLIFT 632
Qy 775 NCILNCPVAFUSFSLNLTFTISPEVIFKILLVVPPLPACLNPLLYILFNPHFKEDLVSL 834
Db 633 DFICWAPISFVALSAILNKLITVSNKILLVLFYPLNSCANPFLYAIFTKEFORDFIL 692
Qy 835 -----RKQYVWTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITVDLPSSVPS 887
Db 693 LSKFGICKRQARQYRGQVPPK---NSTDIQVQK-----VTHMRQGLHNM-----E 736
Qy 888 PAYPVTESCHLS 899
Db 737 DVYELIEKSHLT 748

RESULT 12
US-09-461-657B-5
; Sequence 5, Application US/09461657B
; Patent No. 6635445
; GENERAL INFORMATION:
; APPLICANT: IGARASHI, MASAO
; APPLICANT: MINEGISHI, TAKASHI
; APPLICANT: NAKAMURA, KAZUTO
; TITLE OF INVENTION: NUCLEIC ACID MOLECULARS ENCODING HUMAN LUTEINIZING
; TITLE OF INVENTION: HORMONE-HUMAN CHORIONIC GONADOTROPIN RECEPTOR PROTEIN
; TITLE OF INVENTION: AND TRANSFORMANTS THEREOF
; FILE REFERENCE: 70281/41.226C
; CURRENT APPLICATION NUMBER: US/09/461.657B
; CURRENT FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: 07/757,342
; PRIOR FILING DATE: 1991-09-10
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-657B-5

Query Match 12.7%; Score 595; DB 4; Length 764;
Best Local Similarity 21.9%; Pred. No. 6.9e-39;
Matches 213; Conservative 148; Mismatches 301; Indels 310; Gaps 31;
Qy 9 LLSLPVLIQL-----ATGSSPRSGVLLRGCPHCHCEPDGEMLLRVDCSLGLSELPSN 63
Db 6 LQLVLLLDLPRDLGGMGCSPP-----PCEHQEED----FRVTKDOI----- 44
Qy 64 LSVFTSYLDLSMNISQLPPLPSLRFLLELRAGNALTYIPKGAFTGLYSLKVLMLQN 123
Db 45 -----QRPSLPP----- 52
Qy 124 NQLRHVPTEALQNLRSLSQSLRDANHSYVPPSPCSGLHSRLHLWLDNALTEIPVQAFR 183
Db 53 -----STQTLKLIETHLRTIFSHAFSLNPISRIVS-----IDVTLQOLE 93
Qy 184 SLSLQAMTLALNKIHHIPDVAFGNLSLVVLHLNHNRIHSLGKKCPGLHSLETLDINY 243
Db 94 S-----HSFYNLSKVTHIEIRNTR-----NL 114
Qy 244 NNLDEFPTAIRLSNLKGLHSHNNIRSIPEKAFVGNPSLTIHPYDNPIDQVGRSAFOH 303
Db 115 TYID--PDALKELPLKFLGINTGLKMPFDLTKV-----YSTDIFFI----- 155
Qy 304 LPELRTLTLNGASQITEFPDLTGTANLESRLTGAQISSLP-----QTVCNQLPNLQVLDL 359
Db 156 -----LEITDNPYMT-----SIPWNAFQGLCNETLTLKLYNN 187
Qy 360 SYNLEDLPSPVCOCKLOKIDLRHNEIYEIKVDTFQQLLSRLSNLANKKAIITHPNAFS 419
Db 188 QFTSVQGY-AFN-GTKLDAVYLNKNKY-----LTVIYKDAFG 222

Qy 420 TL-----PSLIKLDSSNLLSFPITGLHGLTHLKLITGNHALQSLISSENFPPELKVIMPYA 476
Db 223 GVYSGPSL--LDVSTQSVTALPSKGLHLEKELIARNTWTKKPLSLSLFLHLTRADLSYP 280
Qy 477 YQCCAFG-----VCENAYKISQWKNQKDNSSMDLHK-----KDA 511
Db 281 SHCCAFKNQKIRGILESLEMCNNESSMQSLRQKRSVNALNSPLHQEYBENLGDGSIIVGYKEK 340
Qy 512 GMFOAQDE-----RDLEDFLDF-----EEDLKALHS-----VOCSP 543
Db 341 SKFODTNHAYVFFEEQDEIIGFQELKNPOEETLQAPDSHYDTICODSEDMVCTP 400
Qy 544 SPGPFKCEHLGLIRIGWTIAVLALCNALVTSTVFRSPLYISPIKLLIGVIAAVN 603
Db 401 KSDEFNFCEDIMGYKFLRIVVFVSLALLGNVFLVLLILTSHYKLVNVPFLMCNLAFAF 460
Qy 604 MLTGSSAVLAGVDATFTGFSFARHGAWENGVCVHVGFLSIFASESSVFLTLTAALBERG 663
Db 461 FCMGYMLLIASVDLYTHSEYINHAIDWQTGPGCNTAGFTVFASLSVYTLTITLERW 520
Qy 664 PSVKYSAKFETKAPFSSLKVII-----LLCALLALMTAAVPLLGGSKYGASPLCLPLPGE 719
Db 521 YAITFAMRLDRKMLRHACALWGGWCCFLAL-----LPLVGLSSAKVSIICLPMDETET 576
Qy 720 PSTGMVVALILLNSLCLFMMTIATYKLYCNL-----DKGLENIDWCSMVKHIALLLFT 774
Db 577 PLALAYIVFVLINIAVAVIVCCCVKIYITVRNPQYNGDK-----DTKIARMAVLIFT 632
Qy 775 NCILNCPVAFUSFSLNLTFTISPEVIFKILLVVPPLPACLNPLLYILFNPHFKEDLVSL 834
Db 633 DFICWAPISFVALSAILNKLITVSNKILLVLFYPLNSCANPFLYAIFTKEFORDFIL 692
Qy 835 -----RKQYVWTRSKHPSLMSINSDDVEKQSCDSTQALVTTSSITVDLPSSVPS 887
Db 693 LSKFGICKRQARQYRGQVPPK---NSTDIQVQK-----VTHMRQGLHNM-----E 736
Qy 888 PAYPVTESCHLS 899
Db 737 DVYELIEKSHLT 748

RESULT 13
US-09-826-509-587
; Sequence 587, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; TITLE OF INVENTION: Protein-Coupled Receptors
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: Patent In Version 2.1
; SEQ ID NO 587
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-587

Query Match 12.7%; Score 595; DB 4; Length 1181;
Best Local Similarity 22.0%; Pred. No. 1.3e-38;
Matches 218; Conservative 145; Mismatches 303; Indels 324; Gaps 32;
Qy 9 LLSLPVLIQL-----ATGSSPRSGVLLRGCPHCHCEPDGEMLLRVDCSLGLSELPSN 63
Db 6 LQLVLLLDLPRDLGGMGCSPP-----PCEHQEED----FRVTKDOI----- 44


```
QY 64 LSVFTSYLDLSMNNISQLLPNPLSLRFLLELRLAGNALTYPKGAFGLYSKVLMLQN 123
Db 45 -----QRIPLSP----- 52
QY 124 NQLRHVPTALQNLRSQSLRDANHSIVYPPSCFGLSLRHLWLDNALTETIPVQAFR 183
Db 53 -----STQTLKLIETHLRTIPSHAFSLNPNISRIYV-----IDVTLQOLE 93
QY 184 SLSALQAMTALANKIHHPDYAFGNLSLVVLHLNHNRIHSLGKCFDGLHSLETLDNY 243
Db 94 S-----HSFYNLKSVTHIEIRNTR-----NL 114
QY 244 NNLDFFPTAIRTLSNKLGFHNSNIRSIPEKAFVGNPSLTIHFDNPIQFVGRSAFQH 303
Db 115 TYID--PDALKELPLKSLAFNTGLKMPFDLTKV-----YSTDIFFI----- 155
QY 304 LPELRTLINGASQITTEFPDLTGTANLESLLTGTQAQISSLP-----QTCNQLPNLQVLDL 359
Db 156 -----LEITDNPYMT-----SIPVNAFQGLCNETLTKLYNN 187
QY 360 SYNLEDLPSFVSCOKLQKIDLRHNEIYEIKVDTFQQLLSLRSLNLANWKIAIHPNAPS 419
Db 188 GFTSVQGYFECT--KLDAVYLNKKY-----LTVIDKDAFG 222
QY 420 TL---PSLIKLDLSSNLLSPFITGLHGLTHLKLGTGNHALQSLSISSENFPELKVEMPYA 476
Db 223 GYVSGPSL--LDVSGTSLVTPSKGLEHLKELIARNSTLTKLALSLSLFLHLTRADLSYP 280
QY 477 YQCCAF-----GVCE-----NAYKISN-OMNKGDNSSMDLHK-----KDA 511
Db 281 SHCCAFKQKIRGILESLMCNESSIETLRQRKSYNALNSPLHQBEBENLGDLSIVGYKEK 340
QY 512 GMFOAQDE-----RLEDPLDF-----BEDLKALHS-----VOCSP 543
Db 341 SKFQDTHNNAHYVFFREQDEIIGFQBELKNPQETLOAPDQSHDYDTICGSEDMVCTP 400
QY 544 SPGPFPKCEHLDDGLRIGVWTIAVLATCNALVTSTVFRSPLYSPIKLLIGVIAAVN 603
Db 401 KSDEFNPECEDIMGYKFLRIWVFWFVSLALLGNVFLVLLITLTSYKLVNVPFLMCLNAPAD 460
QY 604 MLTGSSAVLAGVADFTGSPARHCAWENGVCVHGFSLFSESSVFLTLTAALEBG 663
Db 461 FCMGYLLLIASVDLYTHSEYNNHAIQWQGPCNTAGFTVFASELSVYTLTVITLERW 520
QY 664 FSVKYSAKFETKAPFSSLVII-----LLCALALTAAMAVPLILGGSKYGASPLCLPLPGE 719
Db 521 YAITFAMALDRKIRLRHACAMVGGWVCCFLAL-----LPLVGISSYAKVSICLPMDTET 576
QY 720 PSTMGYVVALILNLSCLFMMTIATYKLYCNL-----DKGDLNIIWDCSMVXKIALLLFTN 775
Db 577 PLALAYIVFVLTNIVAFVIVCCVVKIYITVRNPHNPGDK-----DTKIAKRAMVLIIFTD 632
QY 776 CILNCPVAPLFSSSLINITFTSPEVIFKILLVVVPLACLPLLYLILNPHPKEDLVSL- 834
Db 633 FTCMAPISFYAVSALINKPLITVSNKILLVLYFPINSCANPFLYAIETKAFQDVFILL 692
QY 835 -----RQTYVTWTSKHPSLMSI-----NSDDV-----EKQS 861
Db 693 SKFGICKKQAOAIRGQVRPPKNSTDIQVKVTHMRQGLHNMEDVYELIENSHLTPKKQ 752
QY 862 CDSQALVTFTSSSITYD-----LPPSSVPS 887
Db 753 QISEEYMQTVLAISAEPHHTGLVDSPSSVPS 782
```

RESULT 14

```
US-07-741-453A-59
; Sequence 59, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
```

```
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/741,453A
; FILING DATE: 19911015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 764 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Protein
; US-07-741-453A-59
```

Query Match 12.6%; Score 594; DB 3; Length 764;

Best Local Similarity 21.9%; Pred. No. 8.3e-39;

Matches 213; Conservative 149; Mismatches 299; Indels 312; Gaps 31;

```
QY 9 LLSLPVLQL-----ATGSSPSRSGVLLRGCTHCHCEPDGRMLLRVDCSDGLSELPSN 63
Db 6 LLQLVLLIDLPRDLGGMGSSP-----PCEHQEED---PRVTKCDI----- 44
QY 64 LSVFTSYLDLSMNNISQLLPNPLSLRFLLELRLAGNALTYPKGAFGLYSKVLMLQN 123
Db 45 -----QRIPLSP----- 52
QY 124 NQLRHVPTALQNLRSQSLRDANHSIVYPPSCFGLSLRHLWLDNALTETIPVQAFR 183
Db 53 -----STQTLKLIETHLRTIPSHAFSLNPNISRIYV----- 83
QY 184 SLSALQAMTALANKIHHPDYAFGNLSLVVLHLNHNRIHSLGKCFDGLHSLETLDNY 243
Db 84 -----SIDTLQOLE---SHSFYNLKSVTHIEIRNTR-----NL 114
QY 244 NNLDFFPTAIRTLSNKLGFHNSNIRSIPEKAFVGNPSLTIHFDNPIQFVGRSAFQH 303
Db 115 TYID--PDALKELPLKSLAFNTGLKMPFDLTKV-----YSTDIFFI----- 155
QY 304 LPELRTLINGASQITTEFPDLTGTANLESLLTGTQAQISSLP-----QTCNQLPNLQVLDL 359
Db 156 -----LEITDNPYMT-----SIPVNAFQGLCNETLTKLYNN 187
QY 360 SYNLEDLPSFVSCOKLQKIDLRHNEIYEIKVDTFQQLLSLRSLNLANWKIAIHPNAPS 419
Db 188 GFTSVQGY-APN-GTKLDAVYLNKKY-----SIPVNAFQGLCNETLTKLYNN 187
QY 420 TL---PSLIKLDLSSNLLSPFITGLHGLTHLKLGTGNHALQSLSISSENFPELKVEMPYA 476
Db 223 GYVSGPSL--LDVSGTSLVTPSKGLEHLKELIARNSTLTKLALSLSLFLHLTRADLSYP 280
```



```
Qy 477 YQCCAFG-----VCENAYKISQWNGDSSMDLHK-----KDA 511
Db 281 SHCAFKKQKIRGILBSLNCSSMQSLRQKSVNALNSPLHQEYENLGDSTVGVYKEK 340
Qy 512 GMFOAQDE-----RDLEDFLDF-----EEDLKALHS-----VOCSP 543
Db 341 SKFQDTHNNAHYVFFEEQDEIIIGFQELKNPQETLQAFDSDHYDVTICGSDSDMVCTP 400
Qy 544 SPGPFKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIAAVN 603
Db 401 KSDEFNCPEDIMGYKFLRIVVWFVSLALLGNVFLILLTSHYKLANVPFLMCLNAPAD 460
Qy 604 MLTGSSAVLAGVDAFTGSGFARHGAWENGVGCHVIGFLSIFASESSVFLLLTAAALRG 663
Db 461 FCMGYMLLIASVDLYTHSEYHNAIDWQTGPGCNTAGFTVFASELSVYTLVTITLERW 520
Qy 664 FSVKYSKAFETKAPFSSLUKVI-----LLCALLATMAAVPLLGSGSKYGASPLCLPLPGE 719
Db 521 YAITFAMRLDRKIRLRHACAIMVGWCCFLAL-----LPLVGISSYAKVSICLPMDET 576
Qy 720 PSTWGYWVALILLNSCLFMMTIATYKLYCNL-----DKGLENIWDCSMVKHIALLLFT 774
Db 577 PLALAYIVFVLTINIAFVIVCCCVKIYITVRNPQYNPGDK-----DTKIARMAVLIFT 632
Qy 775 NCILNCPVAFLSFSSLINLTISPVEIKFILLVVPPLPACLNPLLYILFNPHFKEDLVSL 834
Db 633 DFICWAPISFVALSAILNKPLITVNSKILLVLPPLNSCANPFLYAITKAFORDVIL 692
Qy 835 -----RKQTYVWTRSKHPSLMSINSDVKEQSCDSTQALVTFTSSSITYDLPPS-SVP 886
Db 693 LSKFGICKRQAQAYRGQVPPK---NSTDIQVQ-----KVTHDMRQGLHNM 735
Qy 887 SPAYPVTESCHLS 899
Db 736 BDVYELIENSHLT 748

RESULT 15
US-09-826-509-403
; Sequence 403, Application US/09826509
; Patent No. 6806054
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. 6806054-Endogenous, Constitutively Activated Known G
; FILE OF INVENTION: Protein-Coupled Receptors
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/195,747
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 403
; LENGTH: 764
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-826-509-403
```

Query Match 12.6%; Score 593; DB 4; Length 764;

Best Local Similarity 21.9%; Pred. No. 1e-38;

Matches 213; Conservative 148; Mismatches 301; Indels 310; Gaps 31;

```
Qy 9 LLSLPLVLQL-----ATGSSPSRSGVLLRGCPTHCCEPDGRMLLRVDCSLGLSELPSN 63
Db 6 LLQLVLLLDLPRDLGGWCSSP-----PCEHQEED-----FRVTCDOI-----44
```

```
Qy 64 LSVFTSYLDLSMNISQLPNPLPSLRFLEBLRAGNALTYIPKGAFTGLYSKLVMLQN 123
```

```
Db 45 -----ORIPSLPP-----52
Qy 124 NQLHVPTALQNLRSLSQSLRLDANHLISYPPSCFGLSHSLRLWLDNALTEIPVQAFR 183
Db 53 -----STQTLKLIETHLRTIPSHAFSNLPMISRIYVS-----IDVTLQOLE 93
Qy 184 SLSALQAMTLALNKIHHIPDYAFGNLSLVLHLHNNRIHSLGKKCFDGLSHSLTDLNY 243
Db 94 S-----HSFYNSLSKVTHIERNTR-----NL 114
Qy 244 NNLDEFTAIRTNLNLKELGFHSHNNIRSIPEKAFVGNPSLTIHFYNDPQFQVRSFAQH 303
Db 115 TYID--PDALKELPLKFLGIFNTGLKMFDPDLTKV-----YSTDIPFI-----155
Qy 304 LPELRTITLNGASQITEFPDITGTANLESLLTGAQISSLP-----QTVCNQLPNLOVLDL 359
Db 156 -----LEITDNPMYMT-----SIPVNAFQGLCNETLTKLYNN 187
Qy 360 SYNLEDLPSFVQCQKQKIDLRHNEIYEIKVDTFOQLLSRLSLNLANWKIAIHPNAPS 419
Db 188 GFTSVQGY-AFN-GTKLDVAVLNKGY-----LTVIDKDAFG 222
Qy 420 TL-----PSLIKLDLSSNLLSPFITGLHGLTHLKTGNHALQSLISSNFPBLKVIEMPYA 476
Db 223 GVYSGPSL--LDVTSQTSVTALPSKGLHLEKELIARNTWTLLKPLSLSLFHLTRADLSYP 280
Qy 477 YQCCAFG-----VCENAYKISQWNGDSSMDLHK-----KDA 511
Db 281 SHCAFKKQKIRGILBSLNCSSMQSLRQKSVNALNSPLHQEYENLGDSTVGVYKEK 340
Qy 512 GMFOAQDE-----RDLEDFLDF-----EEDLKALHS-----VOCSP 543
Db 341 SKFQDTHNNAHYVFFEEQDEIIIGFQELKNPQETLQAFDSDHYDVTICGSDSDMVCTP 400
Qy 544 SPGPFKCEHLLDGLWIRIGVWTIAVLATCNALVTSTVFRSPLYISPIKLLIGVIAAVN 603
Db 401 KSDEFNCPEDIMGYKFLRIVVWFVSLALLGNVFLILLTSHYKLANVPFLMCLNAPAD 460
Qy 604 MLTGSSAVLAGVDAFTGSGFARHGAWENGVGCHVIGFLSIFASESSVFLLLTAAALRG 663
Db 461 FCMGYMLLIASVDLYTHSEYHNAIDWQTGPGCNTAGFTVFASELSVYTLVTITLERW 520
Qy 664 FSVKYSKAFETKAPFSSLUKVI-----LLCALLATMAAVPLLGSGSKYGASPLCLPLPGE 719
Db 521 YAITFAMRLDRKIRLRHACAIMVGWCCFLAL-----LPLVGISSYAKVSICLPMDET 576
Qy 720 PSTWGYWVALILLNSCLFMMTIATYKLYCNL-----DKGLENIWDCSMVKHIALLLFT 774
Db 577 PLALAYIVFVLTINIAFVIVCCCVKIYITVRNPQYNPGDK-----DTKIARMAVLIFT 632
Qy 775 NCILNCPVAFLSFSSLINLTISPVEIKFILLVVPPLPACLNPLLYILFNPHFKEDLVSL 834
Db 633 DFICWAPISFVALSAILNKPLITVNSKILLVLPPLNSCANPFLYAITKAFORDVIL 692
Qy 835 -----RKQTYVWTRSKHPSLMSINSDVKEQSCDSTQALVTFTSSSITYDLPPSVPSS 887
Db 693 LSKFGICKRQAQAYRGQVPPK---NSTDIQVQ-----VTHMRQGLHNM-----E 736
Qy 888 PAYPVTESCHLS 899
Db 737 DVYELIENSHLT 748
```

Search completed: July 12, 2005, 07:49:29

Job time : 45 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 12, 2005, 21:29:25 ; Search time 197.5 Seconds
(without alignments)
10668.712 Million cell\updates/sec

Title: US-10-751-736-21

Perfect score: 907

Sequence: 1 atggacacctccggctcgg.....catttgtcccatgtctctaa 2724

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3967864

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO spool/US10751736/runat 12072005.084327.23907/app query.fasta_1.2887
-DB=A Genesecp 16Dec04 -QFMT=fastan -SURFIX=oli.rag -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751736 @CGN 1.1.279 @runat 12072005.084327.23907 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

A Genesecp 16Dec04:*
1: Genesecp1980s:*
2: Genesecp1990s:*
3: Genesecp2000s:*
4: Genesecp2001s:*
5: Genesecp2002s:*
6: Genesecp2003as:*
7: Genesecp2003bs:*
8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	2	Aaw93889 Human HG3
2	907	100.0	907	3	Aay90682 Human G p
3	907	100.0	907	6	Abp81968 Human G p
4	907	100.0	907	6	Abp06467 Human G p
5	907	100.0	907	7	Adc22783 Human G p
6	907	100.0	907	7	Adc59150 Human G p
7	907	100.0	907	7	Adc59153 Human Pro
8	907	100.0	907	7	Adg42628 Human G p
9	907	100.0	907	7	Adh14256 Human HG3
10	907	100.0	907	7	Adn40013 Cancer/an

11	907	100.0	907	7	ADN39531	Adn39531	Cancer/an
12	907	100.0	907	7	ADN39628	Adn39628	Cancer/an
13	907	100.0	907	8	ADO29408	Ado29408	Human GPC
14	907	100.0	907	8	ADQ80369	Adq80369	G protein
15	907	100.0	907	8	ADR67868	Adr67868	Human HG3
16	806	88.9	907	3	Aay90687	Aay90687	Human mut
17	806	88.9	907	3	ADC22797	Adc22797	Human G p
18	806	88.9	907	7	ADH14270	Adh14270	Mutated h
19	806	88.9	1145	7	ADF70480	Adf70480	Orphan re
20	716	78.9	907	7	ADG42629	Adg42629	Human G p
21	699	76.0	692	2	AAW93890	Aaw93890	Human HG3
22	621	68.5	883	7	ADB80464	Adb80464	Ovarian c
23	621	68.5	883	7	ADN40012	Adn40012	Cancer/an
24	621	68.5	883	7	ADN39166	Adn39166	Cancer/an
25	592	65.3	693	2	AAY42169	Aay42169	Human LGR
26	202	22.3	202	2	AAW53573	Aaw53573	Human gon
27	197	21.7	282	5	ABP42000	Abp42000	Human ova
28	51	5.6	907	8	ADG75449	Adg75449	Mouse ort
29	51	5.6	907	8	ADO29409	Ado29409	Mouse GPC
30	43	4.7	43	2	AAW93892	Aaw93892	Human HG3
31	43	4.7	43	2	AAW93894	Aaw93894	Human HG3
32	43	4.7	43	2	AAW93893	Aaw93893	Human HG3
33	43	4.7	43	2	AAW93895	Aaw93895	Human HG3
34	43	4.7	43	2	AAW93896	Aaw93896	Human HG3
35	43	4.7	43	2	AAW93891	Aaw93891	Human HG3
36	20	2.2	20	6	ABP82972	Abp82972	G protein
37	20	2.2	20	6	ABP82976	Abp82976	G protein
38	18	2.0	91	2	AAW53577	Aay53577	Human gon
39	18	2.0	244	8	ADO66490	Ado66490	Novel hum
40	18	2.0	266	3	AAH41583	Aah41583	Human ORF
41	18	2.0	736	5	ADG66140	Adg66140	Human LGR
42	18	2.0	736	7	ADN02246	Adn02246	Human par
43	18	2.0	736	8	ADK19411	Adk19411	Human LGR
44	18	2.0	794	5	ADH50810	Adh50810	Human G-p
45	18	2.0	823	5	ADH50809	Adh50809	Human G-p

ALIGNMENTS

RESULT 1

AAW93889

ID AAW93889 standard; protein; 907 AA.

XX

AC AAW93889;

XX

DT 25-JUN-1999 (first entry)

XX

DE Human HG38 protein.

XX

KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain; endocrine system; skeletal muscle; spinal cord; placenta; development; receptor activity modulator.

XX

OS Homo sapiens.

XX

FN WO9915660-A1.

XX

PD 01-APR-1999.

XX

PF 24-SEP-1998; 98WO-US019979.

XX

PR 24-SEP-1997; 97US-0059863P.

XX

(MERI) MERCK & CO INC.

PI Liu Q, Bailey WJ, McDonald TP;

XX

DR WPI; 1999-254711/21.

XX

N-PSDB; AAX23980.

XX Human G-protein coupled glycoprotein hormone receptor HG38.

PS Claim 1a; Fig 2; 74pp; English.

XX This invention describes a novel human G-protein coupled glycoprotein
CC hormone receptor, HG38. Glycoprotein hormone receptors are important in
CC the endocrine system and HG38 may be involved in development and function
CC of the skeletal muscle, spinal cord, placenta and to a lesser extent, the
CC brain. The transgenic animal may be useful for studying tissue and
CC temporal specific expression or activity of the HG38 receptor, as well as
CC for studying the ability of a variety of compounds to act as modulators
CC of HG38 receptor activity
XX

SQ Sequence 907 AA;

Alignment Scores:
Pred. No.: 0 Length: 907
Score: 907.00 Matches: 907
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x AAW93889 (1-907)

QY	1	ATGGACACCTCCCGCTCGGTGCTCTCTGCTCTCTGCTCTGCTGCGGAC	60
DB	1	MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr	20
QY	61	GGGGCAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGTGCCCCACACACTGTCTATGC	120
DB	21	GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCys	40
QY	121	GAGCCGAGCGGAGGATGTTCTCAGGTGGAGTCTCCGACTCGGGCTCTCGAGCTG	180
DB	41	GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu	60
QY	181	CCTTCCAACTCAGGCTCTTCACTCTCTACCTAGACCTCAGTATGAACAACTCAGTCAG	240
DB	61	ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln	80
QY	241	CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTGAGGAGTTACGTTCCGGGAAAC	300
DB	81	LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn	100
QY	301	GCTCTGACATATCTCCAGGAGCATCTACGTCCTTACAGTCTTAAAGTCTTATG	360
DB	101	AlaLeuThrTyrIleProIysGlyAlaPheThrGlyLeuTyrSerLeuLeuMet	120
QY	361	CTGCAGATAATCAGCTAAGACACGTACCACAGAGCTCTGCAGAAATTTGCGAAGCCTT	420
DB	121	LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu	140
QY	421	CAATCCCTCGCTGTGATGCTAACCATCAGTATGTGCCCTCCCAAGCTGTTTCAGTGGC	480
DB	141	GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly	160
QY	481	CTGCATTCCTCGGACCTGTGGCTGATGATGATGATGATGATGATGATGATGATGATG	540
DB	161	LeuHisSerLeuArgHisIleTrpLeuAspAsnAlaLeuThrGluIleProValGln	180
QY	541	GCTTTTGAAGTTTATCGGCATTCGCAAGCATGACCTTGGCCCTCGAACAAATATACACC	600
DB	181	AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnIleHis	200
QY	601	ATACACGATATGCTTTTGAACCTCTCAGCTTGGTAGTGTCTACATCTCCATACAAAT	660
DB	201	IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn	220
QY	661	AGNATCCATCTCTGGGAAGAAATGCTTTGTATGGCTCCAGCTCTAGAGCTTTAGAT	720
DB	221	ArgIleHisSerLeuGlyIysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp	240
QY	721	TTAAATTTACAATAACCTTGAATTTCCCACTTCGCAATTAGGACACTCTCCAACTTAA	780
DB	241	LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuIys	260

QY	781	GAACTAGGATTTTCATAGCAACAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC	840
DB	261	GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluIysAlaPheValGlyAsn	280
QY	841	CCTTCTCTTATTACATATCATTCTTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT	900
DB	281	ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla	300
QY	901	TTTCAACATTTTACCTGAACCTAAGAACACTGACTGTAATGCTGCTCAAAATTAATCTGAA	960
DB	301	PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu	320
QY	961	TTTCTGATTTAACTGGAACTGCAAACTGAGAGTCTGACTTTAACTGGAGACACAGATC	1020
DB	321	PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuGlyAlaGlnIle	340
QY	1021	TCATCTCTCTCAAAACCGTCTGCAATCAGTTACCTATCTCCAAGTCTAGATCTGCT	1080
DB	341	SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer	360
QY	1081	TACAACCTATTAGAGATTTACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAATTCAC	1140
DB	361	TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp	380
QY	1141	CTAAGACATAATGAAATCTACGAAATTAAGTTTGCACACTTTCACAGCTTGTCTAGCCTC	1200
DB	381	LeuArgHisAsnGluIleTyrGluLeuLysValAspThrPheGlnGlnLeuLeuSerLeu	400
QY	1201	CGATCGCTGAATTTGGCTTGGAAACAAAATTCCTATTATTACCCCAATGCAATTTCCACT	1260
DB	401	ArgSerLeuAsnLeuAlaTyrAsnLysIleAlaIleIleHisProAsnAlaPheSerThr	420
QY	1261	TTCATCTCTTAAAGCTGACCTATCTGCTCCAACTCTCTGCTGCTTTCCTTAACT	1320
DB	421	LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr	440
QY	1321	GGTGTACATGTTTAACTCCTTAAATTAACAGGAAATCATGCTCTTACAGAGCTTGATA	1380
DB	441	GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeu	460
QY	1381	TCATCTGAAAATTTCCAGAACTCAAGGTATAGAAATGCCTTATGCTTACAGTGTGT	1440
DB	461	SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCys	480
QY	1441	GCATTTGAGTGTGTGAGAAATGCCTTAAAGATTTCTAATCAATGGAATTAAGTGCACAC	1500
DB	481	AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn	500
QY	1501	AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560
DB	501	SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg	520
QY	1561	GACCTTGAAGATTTCTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG	1620
DB	521	AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln	540
QY	1621	TGTTTCACTTCCCGAGGCCCCCTTCAACCTGTGTGACACCTGCTTGTGCTGCTGCTGATC	1680
DB	541	CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle	560
QY	1681	AGAATTTGAGTGTGTGACCATAGCAGTCTTGGCAGCTTACTTGTAAATGCTTGTGTGCTTCA	1740
DB	561	ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer	580
QY	1741	ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTTAACTGTTGCTGCTGCTGCTG	1800
DB	581	ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla	600
QY	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGGCGTCTGGCTGCTGCTGCTGCTGCTGCTGCT	1860
DB	601	AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr	620

Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Qy 241 CTGCTCCCGAAATCCCTCCCAAGTCTCCGCTTCCTGGAGAGATTACGCTTCGCGGAAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
Qy 301 GCTCTGACATACATTCCTCCAGGGAGCATTCAGTGCCTTTACAGTCTTAAAGTCTTATG 360
Db 101 AlaLeuThrTyrIleProGlyeGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Qy 361 CTGCAGAAATATACCTAAGACACCTACCCACAGAAAGCTCTGCAGAAATTTGGAGAGCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Qy 421 CAATCCCTGCGTCTGGATCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Qy 481 CTGCATTCCTGAGGCACCTGTGGCTGGATGACAAATGGTTAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
Qy 541 GCTTTTAGAAGTTTATCGCATTCGCAAGCCATGACCTTGGCCCTCGAACAAATAATACACCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Qy 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Qy 721 TTAATAATCAATAACTTGATGAATTCCTCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Qy 781 GAACTAGGATTTTCATAGCAACAATATCAGTGCATACCTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Qy 841 CTTCTCTTATTAACAATACATTTCTATGACAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Qy 901 TTTCAACATTTTACCTGAATAAGAACACTGACTCTGAATGGTGGCTCACAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTCTCGATTTTAACCTGGAACCTGCAACCTGGAGAGTCTGACTTTTAACCTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTCTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACCTATTAGAAGATTTACCAGTGTTCAGTCTGCCAAAGCTTCAGAAAAATTCAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTTCAGCAGTGTGCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Qy 1261 TTGCATCCCTAATAAGCTGACCTATCGTCCAACTCTGCTGCTCTTTTCTTCTATAACT 1320

Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Qy 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTTACAGTGCCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Qy 1441 GCATTTGGAGTGTGTGAGAATCCCTATAGAATTTCTTAATCAATGGAATAAAGGTGACAAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTyrAsnLysGlyAspAsn 500
Qy 1501 AGCATATGAGACGACCTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAGATTTCCCTGCTTACTTTGAGGAAGACCTGAAAGCCCTTCATTTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTCACCTTCCCAGGCCCCCTTCAAAACCTGTGAAACACCTGCTGTGATGGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrLeuIle 560
Qy 1681 AGAATGGAGTGTGGACCATAGCAGTCTGCGACTTACTTGTAAATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTTCAGATCCCTCTGACATTTCCCCCATTAACCTGTTAAATGGGGTTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 CGAGTGAACATGCTCACGGGAGTCTCCAGTGCCCGTGGCTGGTGGTGGATGCGTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGCATGTCATT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTGTCCATTTTGTCTTCCAGATCATCTGTTTTCTCTGCTTACTCTGCGACCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Qy 1981 GAGCGTGGGTCTCTGTGAAAATATTCTGCAAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAAGTAAATCATTGCTCTGTGCCCTGTGGCTTGACATGGCCGCGAGTTCCTCCCTG 2100
Db 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGTGGCAGCAAGATGGCGCTCCCTCTCTGCGCTGCTTTCCTCTTTGGGAGGCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCACCATGGCTACATGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTGCTTACCAAGCTCTACTGCAATTTGGCAAGGAGACCTCGAGATATTG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Qy 2281 GACTGCTCTATGGTAAACACATTCGCTGTTGCTTTCACCAACTGATCCTTAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Qy 2341 CCTGTGGCTTCTTGTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCCTGAATA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800

|||||
181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Db
601 ATACGAGACTATGCTTTGGAAACCTCTCCAGCTGGTAGTCTTACATCTCCATAACAAT 660
Qy
201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValLeuHisLeuHisAsnAsn 220
Db
661 AGAATCCACTCCCTGGGAAGAAAGCTTTGATGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy
221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Db
721 TTAATTAACAATAACTGATCAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAAA 780
Qy
241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Db
781 GAATAGAGATTTCAATAGCAACAATATCAGGTGCGATCTGAGAAAGCAATTTGTAGGCAAC 840
Qy
261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Db
841 CCTTCTCTATTACAAATCTTATGACATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Qy
281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Db
901 TTTTCAACATTTACCTGAACATAAGAACACTGACTCTCAATGTGTCCTCACAATAACTGAA 960
Qy
301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Db
961 TTTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Qy
321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Db
1021 TCATCTCTCTCAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
Qy
341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Db
1081 TACAACCTATTAGAAGATTACCCAGTCTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Qy
361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Db
1141 CTAAGACATAATGAAATCTACGAAATTAAGTTGACATTTCCAGCAGTGTCTAGCCTC 1200
Qy
381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Db
1201 CGATCGCTGAATTTGGTGGACAAAATTTGCTATTATTACCCCAATGCATTTTCCACT 1260
Qy
401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Db
1261 TTGGCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCTCGTCTTTTCTTATAACT 1320
Qy
421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Db
1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Qy
441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Db
1381 TCATCTGAAACTTTCCAGACTCAGGTTATAGAAATGCTTATCCCTTACCAGTCTGT 1440
Qy
461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Db
1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAAATAAGGTGACAA 1500
Qy
481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Db
1501 AGCAGTATGAGACCTTCATNAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGACGT 1560
Qy
501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Db
1561 GACCTTGAGATTTCTGCTTGGACTTTGAGGAGACCTGAAAGCCCTTCATTCAAGTGCAG 1620
Qy
521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Db
1621 TGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAAACACCTGCTTGTATGGCTGGCTGATC 1680
Qy

541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrIleLeu 560
Db
1681 AGAATTGAGGTGGACCATAGCAGTCTCTGACACTTACTTCTAATGCTTTGGTGCATTCA 1740
Qy
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Db
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCACTTAACTGTTAATTTGGGGTCAATCGCA 1800
Qy
581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Db
1801 GCAGTGAACATGCTCAGGGAGTCTCAGTCCCGTCTGGCTGGTGTGGATCGCTTCACT 1860
Qy
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Db
1861 TTTGCGACCTTTGACGACATGGTGGCTGGGAGAAATGGGTGGTGGTGCATGCTCAAT 1920
Qy
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Db
1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTCTTCTGCTTACTCTGGCAGCCCTG 1980
Qy
641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
Db
1981 GAGCTGGGTCTCTGTGAAATATCTGCAAAATTTGAAAAGAAAGCTCCATTTCTAGC 2040
Qy
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Db
2041 CTGAAAGTAATCATTTTGTCTCTGCTCTGCTGGCTTGCACATGGCCGCAAGTCCCTCTG 2100
Qy
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Db
2101 CTGGTGGCAGCAGATATGGCCCTCCCTCTCTGCTGCTTCTGCTTGGGGAGGCC 2160
Qy
701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
Db
2161 AGCACCATGGCTACATGCTGCTCTCATCTGCTCAATTCCTTTGCTCTCTCATGATG 2220
Qy
721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Db
2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGAAAGGAGACCTGGAGAAATATTTGG 2280
Qy
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Db
2281 GACTGCTCTATGGTAAAAACACATTCCTGCTTCTTCCAACTGCAAACTGCATCTAAACTGC 2340
Qy
761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Db
2341 CCTGCTGCTTCTGCTCTCTCTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 2400
Qy
781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Db
2401 ATTAAGTTTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Qy
801 IleLysPheIleLeuLeuValValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Db
2461 ATCTTGTTCAACTCCTTAAAGAGGATCTGGTGGCTGAGAAAGCAAACTTACGTC 2520
Qy
821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Db
2521 TGGACAGATCAAAACACCCAGCTGATGTCAAATTAATTAATTAATTAATTAATTAATTA 2580
Qy
841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Db
2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTG 2640
Qy
861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Db
2641 CCTCCAGTTCGCTGCCATCACCAGCTTATCCAGTGCATGAGAGCTGCCATCTTCTCTCT 2700
Qy
881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Db
2701 GTGGCATTTGTCCCATGCTCTC 2721
Qy
901 ValAlaPheValProCysLeu 907
Db


```
QY 1081 TACACCTATTAGAGATTTCACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAAATTCGAC 1140
Db 1141 TAAAGACATAATAGAAATCTAGCAATTAAGAGTTGACACTTTCCAGCAGTCTGCTTAGCCCTC 1200
QY 381 LeuArgHisAsnGluIleTyrgluileysValaspThrPheGlnGlnLeuLeuSerLeu 400
Db 1201 CGATCGCTGAATTTGGCTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnlysiIleAlaIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAAATAAGCTGACCTATCGTCCAACTCTGCTGCTCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerAsnLeuLeuSerPheProIleThr 440
QY 1321 GGGTTACATGGTTTAACTCAGCTTAAATTAACAGGAAATCATGCTTCACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTACAGAGTCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyralaTyrglnCysCys 480
QY 1441 GCATTTGGAGTGTGTGAGAAATGCCTATAGATTTCTCAATCAATGGAATAAAGGTGACAAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrlsIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGGAGCAGCTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAAGANTTCCTGCTGACCTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCCACCTTCCAGGAGCCCTTCAACCCCTGTGAACACCTGCTGTAGTGGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTyrlleuIle 560
QY 1681 AGAATTTGGAGTGTGGACCATAGCAGTCTTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCATTAACCTGTTAAATTTGGGGTCAATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrlleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGCTGGTGGATGCGTTCAC 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGGAGAATGGGGTTGGTGGCATGCTCAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTGTGTCATTTTGTCTTCAGAAATCATCTGTTTTCCTGCTTACTCTGCGACCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCTGGGGTCTCTGTGGAATATTTCTCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrlSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGGCTTTGACCATGGCCGAGTTCCTCCCTG 2100
Db 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAAGATAGGCCCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGAGCCC 2160
Db 701 LeuGlyGlySerLysTyrglyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
QY 2161 AGCACCATGGGCTACATGGTGCCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2220
```

```
Db 721 SerThrMetGlyTyrlMetValAlaLeuLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTCCTACACCAAGCTCTACTGCAATTTTGCAAGGAGAGCTGGAGAAATTTGG 2280
Db 741 ThrIleAlaTyrlThrLysLeuTyrlCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGTAAACACATTGCTGCTGCTTTCACCAACTGCATCCTTAAACTGCG 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGGCTTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCTCTTCTGCTGCTAGTCCACTTCTGCTGCTGCTCAATCCCTTCTCTAC 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyrl 820
QY 2461 ATCTTGTTCATCTCTCACTTTTAAGGAGGATCTGTGTGAGCTTGAGAAAGCAACCTTACGTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrlVal 840
QY 2521 TGGCAAGATCAAAACACCCCAAGCTTGTATGTCAATTAACCTCTGATGTCTGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACCAGTCCAGCATCATTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyrlAspLeu 880
QY 2641 CCTCCCAGTTCCTGCGCCATCACACAGCTTATCCAGTGTAGAGCTGCGCATCTTCTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrlProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTGTGCCCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 5
ADC22783
ID ADC22783 standard; protein; 907 AA.
XX AC ADC22783;
XX DT 18-DEC-2003 (first entry)
XX DE Human G protein-coupled receptor (GPCR) polypeptide #39.
XX KW Human; G protein-coupled receptor; GPCR; transmembrane-6 region; TM6; intracellular-3 region; IC3; receptor.
XX OS Homo sapiens.
XX PN US6555339-B1.
XX PD 29-APR-2003.
XX PF 13-OCT-1998; 98US-00170496.
XX PR 14-APR-1997; 97US-00839449.
XX PR 14-APR-1998; 98US-00060188.
XX PR 26-JUN-1998; 98US-0090783P.
XX PR 07-OCT-1998; 98US-00956677P.
XX PA (AREN-) ARENA PHARM INC.
XX PI Liaw CW, Behan DP, Chalmers DT;
XX WPI; 2003-742861/70.
XX DR N-PSDB; ADC22782.
XX PT Creating a constitutively active version of an endogenous human G protein
```


[illegible]

301 GCTCTGACATATCCCAAGGAGGACATTCACATGGGCTTTACAGTCTTAAAGTCTCTATG 360
101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
361 CTCGAGAATAATCAGCTTAAGACACAGTACCCACAGAGCTCTGACAGATTTCGCAAGCCTT 420
121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
421 CAATCCCTGGCTGTGAGTGTAAACCATCAGCTATGTGTGCCCCCAAGCTGTTTCAGTGGC 480
141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
481 CTCGATTCCTGAGGACCTGCTGGTGAATGCAATGCGTTTACAGAAATCCCGCTCCAG 540
161 LeuHisSerLeuArgHisIleuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
541 GCTTTTAGAAGTTTATCGGATTCGCAAGCCATGACCTTGGCCCTGGAACAAATACACCCAC 600
181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
601 ATACGAGATATGCCCTTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 660
201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
661 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
721 TTAATAATACAACTTGATGAATTCCTCCACTGCAATTAGGACACTCTCCAACTTAAA 780
241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
781 GAATAGGATTTTCATAGCAACAATTCAGTGCATACCTGAGAAAGCATTTGAGCAAC 840
261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
841 CCTTCTCTTATTACATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 900
281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
901 TTTCAACATTTACTGAACACTAAGAACACTGACTCTGAATGGTCCCTCACAATAACTGAA 960
301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
961 TTTCTGATTTAACTGGAACCTGCAACCTGGAGAGTCTGACTTAACTGGAGCACAGATC 1020
321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCAAGTGTAGATCTGTCT 1080
341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
1081 TACAACCTATTAGAAGATTTACCAGTTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1140
361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
1141 CTAAGACATAAATGAATCTACGAAATTTAAAGTTGACACTTTCAGCAGAGTTGCTTAGCCCTC 1200
381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
1201 CGATCGCTGAATTTGGCTGGAAACAAAATTTGCTATTATTATCACCCCAAGTATTTCCACT 1260
401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTGTCGCTTTTCTCTATAACT 1320
421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1380
441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACCAGTCTGT 1440

461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
1441 GCATTTTGGAGTGTGTGAGAAATGCTATAAGATTTCTAATCAATGAATTAAGGTGCAAC 1500
481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
1501 AGCAGTATGACGACCTTCATTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1560
501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
1561 GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
1621 TGTTCACCTTCCCCAGGCGCTTCAAAACCTGTGAAACACCTGCTTGTGATGCTGCTGATC 1680
541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
1681 AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTGAAATGCTTTGGTACTTCA 1740
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCATTTAAACTGTTAATTTGGGTCAATCGCA 1800
581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 1860
1801 GCAGTGAAACATGCTCACGGAGTCTCCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
1861 TTTGGCAGCTTTGCGACGACATGGTGGCTGGTGGGAGAAATGGGTTGGTGGCCATGTCATT 1920
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
1921 GCTTTTCTCTCCATTTTCTGCTTCCAGAAATCATCTGTTTCTGCTGCTGCTGCTGCTGCTGCTG 1980
641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
1981 GAGCGTGGCTTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
2041 CTGAAAGTAAATCATTTTGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
2161 AGCACCATGGCTTACATGCTGCTCTCATCTGCTCAATTCCTTCCCTTCTCTCATGATG 2220
721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
2221 ACCATTGCTTACCAAGCTCTACTGCAATTTGGACAAGGAGACTGGAGATAATTTGG 2280
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
2281 GACTGCTCTATGGTAAACACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
2341 CTGTGGCTTCTTGTGCTTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGTAGTA 2400
781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
2401 ATTAAGTTTATCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
2461 ATCTGTTCAATCCTCACCTTTAAGGAGGATCTGGTGGCTGAGAAAGCAAACTACGCTC 2520

1021	QY	TCATCTCTTCTCTCAAACCGTGCTGCAATCGAGTTCACCTTAATCTCCAAAGTCGTAGATCTGTCT	1080
341	Db	SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer	360
1081	QY	TACAACCTATTAGAAGATTTTACCACGATTTTTCAGTCTGCCCAAAGCTTCACAGAAATTGAC	1140
361	Db	TyrAsnLeuLeuGlnAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp	380
1141	QY	CTAAGACATAATGAAATCTACGAAATTAAGTTGACATTTTCCAGCAGAGTTGCTTAGGCCTC	1200
381	Db	LeuArgHisAsnGlnIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu	400
1201	QY	CGATCGCTGAATTTGGCTTGGACAAATAATGCTATTATTACCCCCCAATGCATTTTCCACT	1260
401	Db	ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr	420
1261	QY	TTGCGATCCCTATATAAGCTGGACCTATCGTCAACCTCTGTCGTCGTCCTTTCTCTATAA	1320
421	Db	LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr	440
1321	QY	GGGTTACATCGTTTAACTCACTTAAAAATTAAACAGGAATCATGCTTTACAGAGCTTGATA	1380
441	Db	GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle	460
1381	QY	TCATCTGAAAACCTTTCAGAACTCAAGGTTATTAGAAATGCTTATGCTTACCAAGTGTGT	1440
461	Db	SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys	480
1441	QY	GCATTTGGAGTGTGAGATGCTATAAGATTTCTAATCAATGGATTAAGATGTCACAC	1500
481	Db	AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn	500
1501	QY	AGCAGTATGCACGACCTTCTATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
501	Db	SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg	520
1561	QY	GACCTTGAAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
521	Db	AspLeuGluAspPheLeuLeuAspPheGluAspLeuLysAlaLeuHisSerValGln	540
1621	QY	TGTTCACTTCCCGACGCCCTTCAACCCCTGTGACACCTGCTGATGCTGCTGGCTGATC	1680
541	Db	CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle	560
1681	QY	AGAATTGGAGTGGACCATAGCAGTCTCGCCACTTCTGTAATGCTTTGGTGACTTCA	1740
561	Db	ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer	580
1741	QY	ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATAAATGTTAAATGGGTGCATCGCA	1800
581	Db	ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla	600
1801	QY	GCAGTGNACATGCTCACGGAGTCTCCAGTCCGCTGCTGGCTGGTGGATGGGTTCACT	1860
601	Db	AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr	620
1861	QY	TTTGGCAGCTTTCACGACATGCTGCTGCTGGGAGAAATGGGTGTGGTGCCTATGTCATT	1920
621	Db	PheGlySerPheAlaArgHisGlyAlaTrpTrpGlnGlyValGlyCysHisValIle	640
1921	QY	GGTTTTTGTGCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980
641	Db	GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu	660
1981	QY	GACGGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCAATTTCTAGC	2040
661	Db	GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer	680
2041	QY	CTGAAAGTAAATCATTTTGTCTGTGCCCTCTGCTGGCCTTGACCATGGCCGAGTTCCTCCTG	2100
681	Db	LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu	700
2101	QY	CTGGGTGGCAGCAAGTATGGCGCTCTCCCTCTCTGCTGCTTTTGCTTTTGGGGAGCCC	2160

[illegible]

Db 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
Qy 121 GAGCCGACGAGAGATGCTCAGGGTGAGACTGCTCCGACCTGGGCTCTCGAGCTG 180
Db 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Qy 181 CCTTCCAACTCAGCGTCTCACCCTCCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 240
Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Qy 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTCTCCCTGGAGAGATTACGCTCTTCGGGAAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuLeuArgLeuAlaGlyAsn 100
Qy 301 GCTCTGACATACATCCCAAGGAGCATTCAGTGCCTTTACGCTCTTAAAGTCTTTATG 360
Db 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Qy 361 CTGCAGATAATCAGCTAAGACACGTACCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Qy 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGCG 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Qy 481 CTGCATCTCCTGAGGCACCTGTGGCTGGATGACAAATGGGTAAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuLeuThrGluIleProValGln 180
Qy 541 GCTTTTAGAAGTTATCGCATTCGCAAGCCATGACCTTGGCCCTCAACAAATAACACCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Qy 601 ATACACGACTATGCTCTTGGAAACCTCTCCAGCTTGGTAGTTCTATCATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisIleAsnAsn 220
Qy 661 AGAATCCATCTCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Qy 721 TTAATAATCAATAAATCTGTAGAAATCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Qy 781 GAACTAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Qy 841 CCTTCTCTATTACATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Qy 901 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTGAATGGTGCCTCAACAATACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTCTGATTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTTCTCAAAACGCTCGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACCTATTAGAAGATTACCCAGTCTTCCAGTCTGCCAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAAGACATAATGAAATCTACGAAATTAAGATTGACATTTCCAGCAGTGTGTTAGCTTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnLeuLeuSerLeu 400

Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCGTATTATTACCCCAATGCATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Qy 1261 TTGCGATCCCTTAATAAAGCTGAGCTATCGTCCAACTCTCTGTCTCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerAsnLeuLeuLeuSerSerPheProIleThr 440
Qy 1321 GGGTTACATGGTTTAACTCATTAAATTAACAGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAAATTTTCAGAACTCAAGCTTATGAAATGCCTTATGCTTACAGTGCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaLysGlnCysCys 480
Qy 1441 GCATTTGGAGTGTGTGAGAATGCCTATAGAATTTCTAATCAATGGAATAAAGGTGACAA 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCATATGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAGCT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAGATTTCTGCTTGTACTTTGAGGAAGACCTGAAAGCCCTTCATTTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTACACCTTCCCGAGGCCCTTCAAACTGTGAACACCTGCTGTGATGGTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATGGAGTGTGACCATAGCAGTTCGCACTTACTTGTATGCTTTTGGTGCATTC 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTACATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGGTCAATCC 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 CACGTGAACATGCTCACGGAGTCTCCAGTCCCGTGTGGTGTGGTGGATGCGTTCAT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGGCAGCTTTGACGACATGCTGCTGGTGGGAGATGGGTGGTGGCCATGCTCAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGCGACCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuLeuThrLeuAlaAlaLeu 660
Qy 1981 GAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGACCATGGCCGCGAGTCCCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGTGGCGCAGCAAGTATGGCGCTCCCTCTCTGCTCCCTGCTTTTGGGGAGGCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCACCATGGGTACATGGTCCGTCTCATCTGTGCTCAATCCCTTTGCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTGCTACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTCGAGAAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760

Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Qy 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisLeuSerTyrValProProSerCysPheSerGly 160
Qy 481 CTGCATTCCCTGAGGCACCTGTGGCTGGATGACAAATGGTTAAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
Qy 541 GCTTTTACAAGTTATCGSCATTGCAAGCATGACCTTGGCCCTGGAACAAATAACACCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHisHis 200
Qy 601 ATACCAGACTATGCGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
Qy 661 AGAATCCACTCCTCGGGAAGAAATGCTTTGATGGGCTCCAGAGCTTAGACACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Qy 721 TTAATATTCAATAACTTGTGATGAATTCCTCCACTGCAATTAGACACTCTCCAACCTTAAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Qy 781 GAACTAGGATTTTCATAGCAACAATATCAGTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Qy 841 CTTCTCTTTATTAACAATATTCATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Qy 901 TTTCAACATTACCTGAACTAGAACACTGACTCTGAATGGTGGCTCCAAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTCTGTGATTTAACTGGAACTGCAAACTGGAGACTCTGACTTTTAACCTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTTCTCAAAACCGTCTGGAATCGATTACTTAATCTCCAAGTCTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACTATTAGAAGATTACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATGCTATTATTATTCACCCCAATGCAATTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Qy 1261 TTGGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCTTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
Qy 1321 GGGTTACATGGTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAAACTTTCCAGAACTCAAGTTATAGAAATGCCCTTATGCTTACCAGTCTGTG 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Qy 1441 GCATTTGGAGTGTGAGAAATGCCTATAAGATTCTTAATCAATGGAATAAAGGTGACAAAC 1500

Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCAGTATGACGACCTTTCATAAGAAAAGATGCTGGAATGTTTCAGGCTCAAGATGACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTCGAAGATTTCCCTGCTTACTTTGAGGAACACCTGAAAGCCCTTCATTTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTTACCTTCCCAGGCCCTTCAAAACCTGTGAAACACCTGTGTGATGGTGTGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTATGCTTTGGTGCATCTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAAATTGGGGTTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCCTGTGGCTGTGTGTGATGCGTTCAT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGGCAGCTTTGCACGACATGGTGGCTGTGGGAGAAATGGGTTGGTGGCCATGTCAAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTCTGCTTACTCTCGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Qy 1981 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGTCCCAATTTCTTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAAGTAAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCACCATGGCTACATGGTCTCATCTCATCTGTGCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTGCTACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Qy 2281 GACTGCTCTATGGTAAAAACACATTGCCCTGTGCTTCTTCCACCAACTGCATCTTAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Qy 2341 CCTGGCTTTCTTGTCTCTCTCTTCTTAAACCTTACATTTATCAGTCTCGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Qy 2401 ATTAAGTTTATCTTCTGCTCTCTCTTCTTAAACCTTACATTTATCAGTCTCGAAGTA 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Qy 2461 ATCTGTTTCAATCTCACTTTAAGGAGGATCTGGTGGCCTCAGAAAGCAAACTACGCTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuLysGlnThrTyrVal 840
Qy 2521 TGGACAAGATCAAAACACCCAGCTTGATGCAATTAATCTCTGTGATGTGCAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860

181 AlapheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
601 ATACCAGACTATGCCCTTTGGAAACCTCCAGCTTGGTAGTTCTACACTCCATAACAAT 660
201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
221 ArgIleHisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
721 TTAATAACAATAACCTTGATGAATTCCTCCACTGCAATTAGACACTCTCCAACTTAAA 780
241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
781 GAACTAGGATTTTCATAGCAACAATATACAGTCGATACCTCGAAGAACGATTTGTAGGCAAC 840
261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
841 CCTTCTCTTATTACATACATTTCTATGACAAATCCATCCATTTGTTGGAGATCTGCT 900
281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
901 TTTCACATTTACCTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAACTGAA 960
301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
961 TTTCTGATTTAACTGGAACCTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuGlyAlaGlnIle 340
1021 TCATCTCTTCTCAAAACGCTGCGAATCAGTTACCTAATCTCCAAGTCTAGATCTGCT 1080
341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
1081 TACAACTATTAGAGAGTTTACCAGTCTCCAGTCTGCCAAAGCTTCAGAAAATTGAC 1140
361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
1141 CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTCTTAGCCTC 1200
381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
1201 CGATCGCTGAATTTGGCTTGGACAAATTTGCTATTATTATCCCAATGCCATTTCCACT 1260
401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
1261 TTGCCATCCCTAATAAAGCTGACCTATCGTCCAACTCCTGCTCTTTTCTTATAACT 1320
421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
1321 GGGTTACATGGTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
1381 TCATCTGAAACTTTCCAGAACTCAAGTTATAGAAATGCTTATGCTTACCAGTCTGT 1440
461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
1441 GCATTTGGAGTGTGTGAGAACTCCCTATAAGATTTCTTAATCAATGAAATAAAGGTGACAAAC 1500
481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
1501 AGCAGTATGAGCAGCTTCATAGAAGAAGTGTGGAATGTTTCAGGCTCAAGATGAACGT 1560
501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
1561 GACCTTGAAGATTTCTGCTTACACTTTGAGGAGACCTGAAAGCCCTTCATTTCACTGCGAG 1620
521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
1621 TGTTCACCTTCCCGAGGCCCTTCAAAACCTGCTGAACACCTGCTGATGGCTGGCTGATC 1680

541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
1681 AGAATTGAGTGTGGACCATAGCAGTTCTGGCAGTTTCTACTTGTANTGCTTTGGTCACTTCA 1740
561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
1741 ACAGTTTTAGATCCCTCTGTACATTTCCCCCATTAACACTGTTAAATGGGGTTCATCGCA 1800
581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
1801 GCACTGAACATGCTCACGGGAGTCTCCAGTGCCCTGCTGGCTGTGTGGATGCGTTCAC 1860
601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
1861 TTTGGCAGCTTTGACACACATGGTGCCTGGTGGGAGAAATGGGTGGTGGCCATGTCATT 1920
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
1921 GGTTTTTTGTCCATTTTGTCTTCAGAATCATCTGTTTCTGCTTACTCTCGCAGCCCTG 1980
641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
1981 GAGCGTGGTTCCTCTGTCAAAATATTCTGCAAAATTTGAAACGAAAGCTTCATTTCTAGC 2040
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
2041 CTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
2101 CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
701 LeuGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
2161 AGCACCATGGCTACATGGTCTCATCTCATCTGCTCAATCCCTGCTGCTGCTGCTGCTGCTG 2220
721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
2221 ACCATTGCTACACCAAGCTCTAGTGCATTTGGACAGGAGAGCTCGAGAAATATTTGG 2280
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
2281 GACTGCTCTATGTGTAAACACATTTGCCCTGTGTGCTTTTCCACCAACTGATCCTAACTGC 2340
761 AspCysSerMetValLysHisIleAlaLeuLeuPheThrAsnCysIleLeuAsnCys 780
2341 CCTGTGGCTTTTGTCTCTCTCTCTTTTATAAACCTTACATTTATCAGTCTCGAAGTA 2400
781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
2401 ATTAAGTTTATCCCTTGTGGTGTAGTCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
2461 ATCTTGTTCATCTCTCACTTTAAGAGGATCTGTGTGAGCCTCGAGAAAGCAAACTACGTC 2520
821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
2521 TGGCAAGATCAAAACACCCAGCTTGATGTCAATTAATCTGTGATGTGTGAAAAACAG 2580
841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
2581 TCCTGTGACTCAACTCAAGCCTTGTAACTTTTACCAGTCCAGCATCATTATCACCTG 2640
861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
2641 CTCCCACTTCCGTGCCATCACAGCTTATCCAGTGTGAGAGCTGAGCTGCTTCTCTCT 2700
881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
2701 GTGGCATTTGCTCCCATGTCTC 2721
901 ValAlaPheValProCysLeu 907

QY 1201 CGATCGCTCAATTTGGCTTGGACAAATTTGCTATTATTACACCCCAATGCTATTTCCACT 1260
DB 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGGCATCCCTAATAAAGCTGGACCTATCGTCAACCTCCTCTGCTGCTTTTCTCTATAACT 1320
DB 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
DB 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAAATTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACACGAGCTGCT 1440
DB 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGGAGTGTGAGAAATGCTATAAGATTTCTAATCAATGGAATAAAGGTGACAA 1500
DB 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGGACACCTTCATAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGACGT 1560
DB 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAGATTTCTGCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTG 1620
DB 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACCTTTCCCGAGCCCTTTCAACCCCTGTGAAACCTGCTGATGCTGGCTGATC 1680
DB 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTTGGAGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1740
DB 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAAACTGTTAATGCGGTCACTCGCA 1800
DB 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGGCTGCTGCTGGTGGTGGATGCTTCACT 1860
DB 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTTGCACACATGCTGCTGGTGGAGAAATGGGGTTGGTGGCCATGTCATT 1920
DB 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTCATTTTTCCTTTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTG 1980
DB 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
QY 1981 GAGCGTGGGTTCCTCTGTGAATAATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAG 2040
DB 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAACTATTTTGTCTGTGCTGCTGCTGGCTTGCATGCGCGAGTTCCCTCG 2100
DB 681 LeuLysValIleIleLeuLeuLysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTGCTGCTGCTGCTGCTTGTGGGAGCCG 2160
DB 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720
QY 2161 AGCACCATGGCTACATGGTCTCATCTTCTCAATTCCTTGTCTTCTCTCATGATG 2220
DB 721 SerThrMetGlyTyrMetValAlaLeuLeuLeuAsnSerLeuLeuMetMet 740
QY 2221 ACCATTGCTTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTGG 2280
DB 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGCTAAACACATTTGCCCTTGTCTCTTCAACCACTGCATCTCAAACTGC 2340

DB 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGGCTTTCTTCT 2400
DB 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCT 2460
DB 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCATCT 2520
DB 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
QY 2521 TGGACAAAGATCAAAACACCAAGCTTGTATCAATTAATCTCTGATGATGTCGAAAAACAG 2580
DB 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2581 TCCTCTGACTCAACTCAAGCCTTGTGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
DB 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CCTCCAGTTCCGTCCTCATCACAGCTTATCCAGTACTGAGAGCTGCCATCTTCTCTCT 2700
DB 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCTCATGCTC 2721
DB 901 ValAlaPheValProCysLeu 907
RESULT 15
ADR67868
ID ADR67868 standard; protein; 907 AA.
XX
AC ADR67868;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human HG38 protein.
XX
KW human; G protein-coupled receptor; GPCR; HG38; colon; lung; cancer.
XX
OS Homo sapiens.
XX
PN WO2004074436-A2.
XX
PD 02-SEP-2004.
XX
PF 11-FEB-2004; 2004WO-US004060.
XX
PR 19-FEB-2003; 2003US-0448959P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Lasek AW;
XX
DR WPI; 2004-652946/63.
DR N-PSDB; ADR67869.
XX
PT Detecting colon or lung cancer, by determining amount of protein in
PT sample, comparing amount of protein to standard, and differential
PT expression of protein in sample indicates colon or lung cancer.
XX
PS Claim 4; SEQ ID NO 1; 79pp; English.
XX
CC This sequence represents the human G protein-coupled receptor (GPCR)
CC known as HG38. The HG38 protein and corresponding nucleic acid, may be
CC used in the method of the invention for detecting colon or lung cancer.
CC The first method involves performing an assay to determine the amount of
CC HG38 in a sample of colon or lung tissue, and comparing the amount of
CC protein to standard, thus detecting expression of protein in sample,
CC where differential expression of protein in sample when compared with the


```
QY 1861 TTTGGCAGCTTTGCACGACATGCTCCCTGGTGGGAGANTGGGTGGTTGCCATGTCTCATTT 1920
Db      |||
621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTCTGCTTACTCTCTGGCAGCCCTG 1980
Db      |||
641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGGTTCTCTGTGAATATTTCTGCAAAATTTGAAAGCTCCATTTCTTAGC 2040
Db      |||
661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCAATTTTCTCTGTGCCCTCTGTGCCCTGTGACATGGCGCAGTTCCTCTG 2100
Db      |||
681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGTGGCAGCAGTATGGCGCTCCCTCTCTGTCCCTGCTTTTGGGGAGGCC 2160
Db      |||
701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPhePheGlyGluPro 720
QY 2161 AGCACCATGGGCTACATGGTCTCTCATTTGCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db      |||
721 SerThrMetGlyTyrMetValAlaAlaLeuLeuLeuAsnSerLeuLeuCysPheLeuMetMet 740
QY 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGAGACCTGGAGAATATTGG 2280
Db      |||
741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGTAAACACATTCCTGCTCTTCCCACTGCTCAATCCCTTCTCTAC 2340
Db      |||
761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGGCTTCTTCTCTCTCTCTCTTAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Db      |||
781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCCTTCTCTGGTGGTAGTCCCACTTCTGCTGCTCAATCCCTTCTCTAC 2460
Db      |||
801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCATCTCCTCACTTTAAGGAGGATCTGGTGAGCTGAGAAAGCAACCTACGTC 2520
Db      |||
821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
QY 2521 TGGACAAGATCAAAACACCCAGCTTGATGCAATTAATCTGTGATGATGTCGAAAAACAG 2580
Db      |||
841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2581 TCCTGTGACTCACTCAAGCCTTGTAACCTTTACAGCTCCAGCATCATTATGACCTG 2640
Db      |||
861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CCTCCAGTCCGTCCTCCATCACAGCTTATCCAGTGACTGAGAGCTGCATCTTCTCTCT 2700
Db      |||
881 ProProSerSerValProSerProLalaIyProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCCATGTCTC 2721
Db      |||
901 ValAlaPheValProCysLeu 907
```

Search completed: July 13, 2005, 03:49:43
Job time : 375.5 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	907	100.0	907	4	US-09-170-496D-264		Sequence 264, App
2	806	88.9	907	4	US-09-170-496D-278		Sequence 278, App
C 3	9	1.0	131	4	US-09-252-991A-17184		Sequence 17184, A
4	9	1.0	135	4	US-09-252-991A-18526		Sequence 18526, A
C 5	9	1.0	268	4	US-09-252-991A-17449		Sequence 17449, A
6	9	1.0	644	2	US-08-866-757-2		Sequence 2, Appli
7	9	1.0	644	3	US-09-153-593-2		Sequence 2, Appli
8	9	1.0	1133	4	US-10-101-464A-809		Sequence 809, App
9	8	0.9	99	4	US-09-482-273-258		Sequence 258, App
10	8	0.9	118	4	US-09-252-991A-25344		Sequence 25344, A
C 11	8	0.9	149	4	US-09-252-991A-24978		Sequence 24978, A
C 12	8	0.9	157	4	US-09-252-991A-28441		Sequence 28441, A

SUMMARIES

[illegible]

US-10-751-736-21 (1-2724) x US-09-170-496D-264 (1-907)

QY	1	ATFGACACTCCCGCTCGGTGCTCTCTGCTTCCTGCTGCTGCTGCTGGGACC	60
Db	1	MetAspThrSerArgLeuGlyValLeuLeuSerLeuProValLeuLeuGlnLeuAlaThr	20
QY	61	GGGGGCAGCTCTCCCAAGTCTGGTGTGTGCTGAGGGGCTGCCCCACACACTGTCAATTGC	120

Db 21 GlyGlySerSerProArgSerGlyValLeuLeuArgGlyCysProThrHisCysHisCys 40
QY 121 GAGCCGCGAGGATGTTGCTCAGGGTGGATGCTCCGACCTCGGGCTCTCGGAGCTG 180
Db 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
QY 181 CTTTCAACCTCAGGCTTACCTCCCTACCTAGACCTCAGTATGAAACAACATCAGTCAG 240
Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
QY 241 CTGCTCCCGAATCCCTCGCAGTCTCGCTTCTCGGAGGAGTACGCTCTTCGGGAAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluLeuArgLeuAlaGlyAsn 100
QY 301 GCTCTGACATACATTTCCCAAGGAGCATTCACCTGGCCCTTTACAGCTTTAAAGTCTTATG 360
Db 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
QY 361 CTGCAGAAATATCAGCTAAGACAGCTACCCACAGAGCTCTGCAGAAATTCGGAAGCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
QY 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
QY 481 CTGCATTCCTGAGCACCTGCTGGTGGATGACAATGCGTTAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAspAsnAlaLeuThrGluIleProValGln 180
QY 541 GCTTTTAGAAGTTATCGCATTCAGCCATGACCTGGCCCTCGAACAATAATACACCAC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
QY 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnAsn 220
QY 661 AGAATCCACTCCTCGGAAAGAAATGCTTTGTATGGGCTCCACAGCTAGAGACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
QY 721 TTAATATTACAATAACCTTGATGAATTCCTCCACTGCAATTAGCACACTCTCCACCTTAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
QY 781 GAATAGGATTTTCATAGCAACAATATCAGTCTGATACCTGAGAAAGCATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
QY 841 CTTCTCTTATTAACATATTCATAGCAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
QY 901 TTTCAACATTTACCTGAACATAGAACACTGCTCTGAATGGTGCCTCACAAATACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
QY 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuGlyAlaGlnIle 340
QY 1021 TCATCTCTTCCAAACCGCTCGCAATCAGTTACCTAATCTCCAAAGTGTAGATCTGCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACCTATTAGAGATTTACCGAGTTTTCAGTCTGCGCAAAAGCTTCAGAAATTCAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTCCAGCAGTGTGCTTACGCTC 1200

Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTCTCTATTATTACCCCAATGCGATTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCGCATCCCTAAATAAGCTCGACTCGTCCCAACCTCCCTGCTGCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGTTTAACTCAGCTTAAATTAACAGAGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAACTTTCCAGAACTCAGGTTATAGAAATGCCTTATGCTTACAGGCTGCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGGAGTGTGTGAGAATGCCTATAAGATTTCTTAATCAATGGAATAAAAGGTGACAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGGACGACCTTCATAGAAGATGCTCGAAATGTTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAGATTTCCCTGCTTACCTTGAGAGAACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluLeuAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACCTCTCCCGAGCCCTTCAACCCCTGTAACACCTGTAACCTGCTGTGAGTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATGGAGTGTGGACCATAGCAGTCTCGACACTTACTGTGAATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTCTACATTTCCCCCATTAACCTGTTAATTTGGGGTCATCCGA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 CGAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGTGGTGTGCGATGCGTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGCGAGCTTTGACACGACATGCTGCTGGTGGAGAAATGGGGTGGTGGTCCATGTCATT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTTGTCCATTTTGTCTCAGATCATCTGTTTTCCTGCTTACTCTGCGAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCGTGGTGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGTGGCAGCAGATGAGCGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGCTCACCAAGCTCTACGTCAATTTGGCAATTTGGAGAGAGCTCGAGATATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760

Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAAATAAGCTGACCTATCGTCCAAACCTCCTGTCGTCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
QY 1321 GGGTTACATGGTTAACTCACCTTAAATTAACAGGAATCATGCCCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAAACCTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAGTGTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGAGTGTGTGAGAATGCCCTATAAGATTCTTAATCAATGAATAAAGGTGACAAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTTCCTGCTTACCTTTCAGGAAGACCTGAAAGCCCTTCATTTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACTTCCCAAGGCCCTTCAAACCCCTGTGAAACACCTGCTTGTAGTGGCTGTATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTGGAGTGTGACCATAGCAGTCTGACACTTCTGTAATGCTTTCGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAAACTGTTAAATGGGTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyriLeSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTCCGCTGCTGGTGGTGGTGGTGGTGGTGGT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTGTGCAGCTTGTGCACGATGCTCCCTGGTGGGAGAATGGTGGTGGTGGTGGTGGTGGT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTGTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCCTGCTTACTCTGCGACCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
QY 1981 GAGCTGGGTTCTCTGTGAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAGTAATCATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGTGGCGACGACATATGGCCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPhePheGlyGluPro 720
QY 2161 AGCACATGGGCTACATGTCGCTCTCATCTGCTCAATTCCTTGGCTTCTTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAATATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuLysCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGT-AAAACATTTGCCCTGTGCTTTCACCAACTGCATCCTAAACTG 2339
Db 761 AspCysSerMetLysLys-HisIleAlaLeuLeuPheThrAsnCysIleLeuAsnCy 780

QY 2340 CCCTGTGGCTTCTTGTGCTTCTCCTCTTTAATAAACCTTACATTTATCAGTCCCTGAAGT 2399
Db 780 sProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVa 800
QY 2400 AATTAAGTTTATCCTTCTTGGTGGTAGTCCCACTTCTCCGCAATGTCTCAATCCCTTCTCTA 2459
Db 800 IileLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTy 820
QY 2460 CATCTTGTCAATCCTCATTAAAGGAGATCTGGTGGAGCTTGAGAAAGCAAACTACGT 2519
Db 820 xIleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVa 840
QY 2520 CTGGACAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTCTGATGATGTCGAAAAACA 2579
Db 840 lTrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2580 GTCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTACCAGCTCCAGCATCACTTATGACCT 2639
Db 860 nSerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLe 880
QY 2640 GCCTCCAGGTTCCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTC 2699
Db 880 uProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSe 900
QY 2700 TGTGCACTTGTGCCATGTCTC 2721
Db 900 rValAlaPheValProCysLeu 907

RESULT 3

US-09-252-991A-17184
; Sequence 17184, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17184
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17184

Alignment Scores:
Pred. No.: 10.8 Length: 131
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.04% Indels: 0
DB: Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-17184 (1-131)

QY 64 CCCCCTGCGCAGCTGCGACGACACAG 38
Db 63 ProArgSerProAlaAlaAlaGln 71

RESULT 4

US-09-252-991A-18526
; Sequence 18526, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18526
LENGTH: 135
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18526

Alignment Scores:
Pred. No.: 10.7 Length: 135
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-18526 (1-135)

QY 56 CGACCGGGGCGACCTCCAGGTCTG 82
DB 116 ArgProGlyAlaLeuProGlyLeu 124

RESULT 5

US-09-252-991A-17449
Sequence 17449, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17449
LENGTH: 268
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-17449

Alignment Scores:
Pred. No.: 9.65 Length: 268
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.04% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-17449 (1-268)
QY 63 CCGGTGCGGCGAGTGCAGGACAGG 37
DB 160 ProGlyArgGlnLeuGlnGlnHisArg 168

RESULT 6

US-08-866-757-2
Sequence 2, Application US/08866757
Patent No. 5858716
GENERAL INFORMATION:
APPLICANT: ELISHOURBAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGSMAN, DEK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)

NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

SEQUENCE CHARACTERISTICS:

LENGTH: 644 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-866-757-2

Alignment Scores:

Pred. No.: 8.43 Length: 644
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x US-08-866-757-2 (1-644)

QY 1420 CCTTATGCTTACCAGTGTGTCATT 1446

DB 158 ProTyAlaTyGlnCysCysAlaPhe 166

RESULT 7

US-09-153-593-2
Sequence 2, Application US/09153593A
Patent No. 6174994
GENERAL INFORMATION:
APPLICANT: ELISHOURBAGY, NABIL A
APPLICANT: LI, XIAOTONG
APPLICANT: BERGSMAN, DEK J
TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
FILE REFERENCE: GH-70055-1
CURRENT APPLICATION NUMBER: US/09/153,593A
CURRENT FILING DATE: 1998-09-15
EARLIER APPLICATION NUMBER: 08/866,757
EARLIER FILING DATE: 1997-05-30
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 644
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-153-593-2

Alignment Scores:
Pred. No.: 8.43 Length: 644
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 3 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-153-593-2 (1-644)

Qy 1420 CCTATGCTTACCAAGTCTGTGCAATTT 1446
Db 158 ProtyrAlaIyrGInCysAlaIaphe 166

RESULT 8

US-10-101-464A-809
; Sequence 809, Application US/10101464A
; Patent No. 6768041
; GENERAL INFORMATION:
; APPLICANT: Strabala, Timothy
; APPLICANT: Nieuwenhuizen, Nicolaas
; APPLICANT: Higgins, Colleen M.
; TITLE OF INVENTION: Compositions Isolated from Plant Cells
; TITLE OF INVENTION: and Their Use in the Modification of Plant Cell Signaling
; FILE REFERENCE: 11000.1020C2
; CURRENT APPLICATION NUMBER: US/10/101,464A
; PRIOR FILING DATE: 2002-03-18
; PRIOR APPLICATION NUMBER: 09/704,302
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 09/228,986
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/162,866
; PRIOR FILING DATE: 1999-11-01
; PRIOR APPLICATION NUMBER: PCT/US00/00724
; PRIOR FILING DATE: 2000-01-11
; NUMBER OF SEQ ID NOS: 989
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 809
; LENGTH: 1133
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-10-101-464A-809

Alignment Scores:
Pred. No.: 7.73 Length: 1133
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-10-101-464A-809 (1-1133)

Qy 1060 CTCGAAGTGTAGATCTGTCTTACAAC 1086
Db 175 LeuGInValLeuAspIeuSerTyrAen 183

RESULT 9

US-09-482-273-258
; Sequence 258, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15

; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-258

Alignment Scores:
Pred. No.: 102 Length: 99
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-482-273-258 (1-99)

Qy 2230 TACACCAAGCTCTACTGCAATTTG 2253
Db 33 TyrThrIysLeuTyrCysAsnLeu 40

RESULT 10

US-09-252-991A-25344
; Sequence 25344, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25344
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25344

Alignment Scores:
Pred. No.: 99.4 Length: 118
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 4 Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-25344 (1-118)

Qy 42 GCTGCTGCAGCTGGCGACCGGGGG 65
Db 33 AlaAlaAlaAlaGlyAspArgGly 40

RESULT 11

US-09-252-991A-24978
; Sequence 24978, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18

: PRIOR APPLICATION NUMBER: US 60/094,190
 : PRIOR FILING DATE: 1998-07-27
 : NUMBER OF SEQ ID NOS: 33142
 : SEQ ID NO 24978
 : LENGTH: 149
 : TYPE: PRT
 : ORGANISM: Pseudomonas aeruginosa
 : US-09-252-991A-24978

Alignment Scores:		
Pred. No.:	95.9	Length: 149
Score:	8.00	Matches: 8
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	0.92%	Indels: 0
DB:	4	Gaps: 0

US-10-751-736-21 (1-2724) x US-09-252-991A-24978 (1-149)

Qy 84 ACCAGACCTGGGAGAGCTGCCCC 61
Db 36 ThrArgProGlyVArgAlaAalaPro 43

RESULT 12

US-09-252-991A-28441
; Sequence 28441, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

```

/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 28441

```

```

; LENGTH: 157
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28441

```

Alignment Scores:					
Pred. No.:	95.1	Length:	157		
Score:	8.00	Matches:	8		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	0.92%	Indels:	0		
DR:	4	Gaps:	0		

US-10-751-736-21 (1-2724) x US-09-252-991A-28441 (1-157)

Qy 78 CCTGGAGAGCTGCCCCCGGTCGC 55
|||
Db 40 ProGlyArgAlaAalaProGlyArg 47

RESULT 13

US-09-252-991A-28201
; Sequence 28201, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

: GENETIC INFORMATION: Rubenfield et al.
 : APPLICANT: Marc J. NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 : TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 : TITLE OF INVENTION:

```

/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/

```

```

; PRIOR FILING DATE: 1998-07-27
;
; NUMBER OF SEQ ID NOS: 33142
;
; SEQ ID NO 28201
;
; LENGTH: 161
;
; TYPE: PRT
;
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28201

```

Alignment Scores:			
Pred. No.:	94.8	Length:	161
Score:	8.00	Matches:	86
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.88%	Indels:	0
DB:	4	Gaps:	0

US-10-751-736-21 (1-2724) x US-09-252-991A-28201 (1-161)

Qy	56	CGACCGGGGCAGCTCTCCAGGT	79
Db	79	ArgProGlyAlaAlaLeuProGly	86

RESULT 14

US-09-270-767-60345
; Sequence 60345, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:

```

; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17

```

```

; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 60345
; LENGTH: 203

```

ORGANISM: Drosophila melanogaster
US-09-270-767-60345

Alignment Scores:				
Pred. No.:	91.5	Length:	203	
Score:	8.00	Matches:	8	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	0.88%	Indels:	0	
DB:	4	Gaps:	0	

US-10-751-736-21 (1-2724) X US-09-270-767-60345 (1-203)

Qy	486 TTCCCTGAGGCACCTGTGGTGA 509
Db	68 PheProGluAlaProValAlaGly 75

RESULT 15

US-09-482-273-167
; Sequence 167, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:

```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267

```


Wed Jul 13 13:55:02 2005

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 167
; LENGTH: 207
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-482-273-167

Alignment Scores:
Pred. No.:      91.2      Length:      207
Score:          8.00      Matches:      8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    0.88%      Indels:      0
DB:             4         Gaps:        0

US-10-751-736-21 (1-2724) x US-09-482-273-167 (1-207)

Qy  2428  CCACCTTCCTGCATGTCTCAATCCC 2451
      |||||||
Db   47  ProLeuProAlaCysLeuAsnPro 54

Search completed: July 13, 2005, 04:00:05
Job time : 75 secs
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model
Run on: July 13, 2005, 03:40:32 ; Search time 194.5 Seconds
(without alignments)
10821.262 Million cell updates/sec

Title: US-10-751-736-21
Perfect score: 907
Sequence: 1 atggacactccgctcg.....cattgtccatgtctctaa 2724

Scoring table:
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1726218 seqs, 386331768 residues
Word size: 1

Total number of hits satisfying chosen parameters: 3311304

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-DB=Published Applications AA -QWTF=fastan -SUFFIX=oli.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MAIRIX=oligo
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10751736 @CGN 1.1 278 @runat 12072005 084330 23973
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :
1: /cgn2_6/ptodata/2/pubaa/US07_PUBCOMB.pcp.*
2: /cgn2_6/ptodata/2/pubaa/PCT_NEW_PUB.pcp.*
3: /cgn2_6/ptodata/2/pubaa/US06_NEW_PUB.pcp.*
4: /cgn2_6/ptodata/2/pubaa/US06_PUBCOMB.pcp.*
5: /cgn2_6/ptodata/2/pubaa/US07_NEW_PUB.pcp.*
6: /cgn2_6/ptodata/2/pubaa/PCTUS_PUBCOMB.pcp.*
7: /cgn2_6/ptodata/2/pubaa/US08_NEW_PUB.pcp.*
8: /cgn2_6/ptodata/2/pubaa/US08_PUBCOMB.pcp.*
9: /cgn2_6/ptodata/2/pubaa/US09A_PUBCOMB.pcp.*
10: /cgn2_6/ptodata/2/pubaa/US09B_PUBCOMB.pcp.*
11: /cgn2_6/ptodata/2/pubaa/US09C_PUBCOMB.pcp.*
12: /cgn2_6/ptodata/2/pubaa/US09_NEW_PUB.pcp.*
13: /cgn2_6/ptodata/2/pubaa/US10A_PUBCOMB.pcp.*
14: /cgn2_6/ptodata/2/pubaa/US10B_PUBCOMB.pcp.*
15: /cgn2_6/ptodata/2/pubaa/US10C_PUBCOMB.pcp.*
16: /cgn2_6/ptodata/2/pubaa/US10D_PUBCOMB.pcp.*
17: /cgn2_6/ptodata/2/pubaa/US10E_PUBCOMB.pcp.*
18: /cgn2_6/ptodata/2/pubaa/US10F_NEW_PUB.pcp.*
19: /cgn2_6/ptodata/2/pubaa/US11A_PUBCOMB.pcp.*
20: /cgn2_6/ptodata/2/pubaa/US11_NEW_PUB.pcp.*
21: /cgn2_6/ptodata/2/pubaa/US60_NEW_PUB.pcp.*
22: /cgn2_6/ptodata/2/pubaa/US60_PUBCOMB.pcp.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	907	100.0	907	10	US-09-965-536A-16	Sequence 16, Appl
2	907	100.0	907	10	US-03-970-944-26	Sequence 26, Appl
3	907	100.0	907	14	US-10-251-385-264	Sequence 264, App
4	907	100.0	907	14	US-10-225-567A-422	Sequence 422, App
5	907	100.0	907	15	US-10-295-027-849	Sequence 849, App
6	907	100.0	907	15	US-10-295-027-946	Sequence 946, App
7	907	100.0	907	15	US-10-295-027-1331	Sequence 1331, Ap
8	907	100.0	907	15	US-10-751-736-84	Sequence 84, Appl
9	907	100.0	907	17	US-10-482-029-158	Sequence 158, App
10	806	88.9	907	14	US-10-251-385-278	Sequence 278, App
11	806	88.9	1145	17	US-10-505-486-103	Sequence 103, App
12	716	78.9	907	10	US-09-970-944-27	Sequence 27, Appl
13	621	68.5	883	15	US-10-295-027-484	Sequence 484, App
14	621	68.5	883	15	US-10-295-027-1330	Sequence 1330, Ap
15	621	68.5	883	15	US-10-173-999-28	Sequence 28, Appl
16	592	65.3	693	17	US-10-851-470-4	Sequence 4, Appli
17	197	21.7	282	15	US-10-264-049-3132	Sequence 3132, Ap
18	51	5.6	907	14	US-10-271-078-10	Sequence 10, Appl
19	20	2.2	20	14	US-10-225-567A-1645	Sequence 1645, Ap
20	20	2.2	20	14	US-10-225-567A-1649	Sequence 1649, Ap
21	18	2.0	736	10	US-03-851-595-8	Sequence 8, Appli
22	18	2.0	736	15	US-10-664-667-8	Sequence 8, Appli
23	18	2.0	794	14	US-10-270-336-7	Sequence 7, Appli
24	18	2.0	823	14	US-10-270-336-6	Sequence 6, Appli
25	18	2.0	828	10	US-09-970-944-25	Sequence 25, Appli
26	18	2.0	828	14	US-10-270-336-2	Sequence 2, Appli
27	18	2.0	828	14	US-10-225-567A-581	Sequence 581, App
28	18	2.0	893	10	US-09-970-944-24	Sequence 24, Appl
29	18	2.0	915	14	US-10-270-336-5	Sequence 5, Appli
30	18	2.0	915	15	US-10-331-496A-68	Sequence 68, Appl
31	18	2.0	915	15	US-10-295-027-1336	Sequence 1336, Ap
32	18	2.0	928	16	US-10-783-528-101	Sequence 101, App
33	18	2.0	940	10	US-09-970-944-6	Sequence 6, Appli
34	18	2.0	967	10	US-03-851-595-2	Sequence 2, Appli
35	18	2.0	967	10	US-03-851-595-11	Sequence 11, Appl
36	18	2.0	967	15	US-10-331-496A-95	Sequence 95, Appl
37	18	2.0	967	15	US-10-684-667-2	Sequence 2, Appli
38	18	2.0	967	15	US-10-684-667-11	Sequence 11, Appl
39	18	2.0	967	16	US-10-684-667-32	Sequence 32, Appl
40	18	2.0	1005	14	US-10-176-847-90	Sequence 90, Appl
41	18	2.0	1205	17	US-10-505-486-64	Sequence 64, Appl
42	18	2.0	1205	17	US-10-505-486-64	Sequence 64, Appl
43	17	1.9	120	14	US-10-276-340-2	Sequence 2, Appli
44	17	1.9	948	16	US-10-328-036-10	Sequence 10, Appl
45	16	1.8	16	14	US-10-225-567A-1646	Sequence 1646, Ap

ALIGNMENTS

RESULT 1
US-09-965-536A-16
; Sequence 16, Application US/09965536A
; Publication No. US20030027323A1
; GENERAL INFORMATION:
; APPLICANT: FEDER, J. N.
; APPLICANT: MINTIER, G.
; APPLICANT: RAMANATHAN, C. S.
; APPLICANT: HAWKEN, D. R.
; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMV5,
; TITLE OF INVENTION: EXPRESSED HIGHLY IN BRAIN AND OVARIAN TISSUES
; FILE REFERENCE: D0041NP
; CURRENT APPLICATION NUMBER: US/09/965,536A
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/235,713
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/261,781
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/306,605
; PRIOR FILING DATE: 2001-07-19

Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
QY 901 TTTCAACATTACTGAACACTAGAACACTGACTCTGAATGGTGGCTCACAAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuAArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
QY 961 TTTCTGATTAACTGGAAGTGAACCTGGAAGTCTGACTTAACTGAGGACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
QY 1021 TCATCTCTCTCAAAACCGCTGCAATCAGTTACCTAAATCTCAAGTGTAGATCTGTCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACCTATTAGAAGATTACCAAGTTTTCAGTCTGCAAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGACATAAATCAATCTACGAATAAAGTTGACACTTCCAGCAGTGTGCTTACGCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
QY 1201 CGATCGCTCAATTTGGCTTGGACAAATTTGCTATTATTACCCCAATGCATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAATAAGCTGACCTATCGTCCAACTCTCTGCTGCTTTTCTCTAATCT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
QY 1321 GGGTTACATGGTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAATTTTCCAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAGTGTCT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTGGAGTGTGAGAAATGCCATAAGATTTCTTAATCAATGAATAAAGTGACAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 ACCAGTATGACGACCTTCATAGAAGATGCTGGAATGTTTCAGGCTCAAGATCAAGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTCTCTGCTTCACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGCGAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACTTCCCGAGGCCCTTCAAAACCCCTGTGAACACCTCTGTGATGGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTGAGTGTGACCATGACGATCTGGGACCTTACTTGTGAATGCTTGTGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCATTAACCTGTAATTTGGGCTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 CGAGTGAACATGCTCACGGAGCTCCAGTGGCGCTGGCTGGTGTGATGCGTTCATCT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTTGGCAGCTTTCAGGACATGCTGCTGGTGGGAGATGGGGTGGTGGCCATGCTCAAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTCTGCTTACTCTGCGACGCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660

QY 1981 GAGCGTGGGTTCTCTGTGAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATATTGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 681 LeuLysValIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
QY 2161 AGCACATGGGCTACATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAGGAGACTGAGAGATATTTCG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGCTAAACACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGGCTTTCT 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCATCTCACTTTAAGGAGGATCTGGTGAGCTGAGAAAGCAAACTACGTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrVal 840
QY 2521 TGGCAAGATCAAAACACCAAGCTTGTGATGCTCAATTAATCTGTGATGCTGCAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
QY 2581 TCCTGTGATCACTCAAGCTTGTGATCACTTACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
QY 2641 CTTCCAGTTCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGCTCCCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 5

US-10-295-027-849
; Sequence 849, Application US/10295027
; Publication No.: US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13

Db 41 GluProAspGlyArgMetLeuLeuArgValAspCysSerAspLeuGlyLeuSerGluLeu 60
Qy 181 CCTTCCAACTCAGGCTCTTCACTCTTACCTAGACCTCAGTATGAAACAAATCAGTCAG 240
Db 61 ProSerAsnLeuSerValPheThrSerTyrLeuAspLeuSerMetAsnAsnIleSerGln 80
Qy 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTTCTGAGAGAGTACGTCCTGGCGGAAC 300
Db 81 LeuLeuProAsnProLeuProSerLeuArgPheLeuGluGluLeuArgLeuValGlyAsn 100
Qy 301 GCTCTGACATACATCCCAAGGAGCATCTACCTGGCCTTTTACAGTCTTAAAGTCTTATG 360
Db 101 AlaLeuThrTyrIleProLysGlyAlaPheThrGlyLeuTyrSerLeuLysValLeuMet 120
Qy 361 CTGCAGATAATACGCTAAGCACACTACCCACAGAGCTCTGCAGAAATTCGGAAGCCCTT 420
Db 121 LeuGlnAsnAsnGlnLeuArgHisValProThrGluAlaLeuGlnAsnLeuArgSerLeu 140
Qy 421 CAATCCCTCGCTCGATGCTAAACACATCAGCTATGCCCCCAAGCTGTTTCAGTGGC 480
Db 141 GlnSerLeuArgLeuAspAlaAsnHisIleSerTyrValProProSerCysPheSerGly 160
Qy 481 CTGCAATCCCTCAGGACACTCTGGCTGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Db 161 LeuHisSerLeuArgHisLeuTrpLeuAspAsnAlaLeuThrGluIleProValGln 180
Qy 541 GCTTTTAGAAGTTATCGGCAATGCAACCATCAGCTATGCCCCCAAGCTGTTTCAGTGGC 600
Db 181 AlaPheArgSerLeuSerAlaLeuGlnAlaMetThrLeuAlaLeuAsnLysIleHis 200
Qy 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTTACATCTCCATAACAAT 660
Db 201 IleProAspTyrAlaPheGlyAsnLeuSerLeuValLeuHisLeuHisAsnAsn 220
Qy 661 AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
Db 221 ArgIleHisSerLeuGlyLysLysCysPheAspGlyLeuHisSerLeuGluThrLeuAsp 240
Qy 721 TTAATTAACAATACCTTGATGAATTCCTCCACTGCAATTAGACACTCTCCAACTTAA 780
Db 241 LeuAsnTyrAsnAsnLeuAspGluPheProThrAlaIleArgThrLeuSerAsnLeuLys 260
Qy 781 GAATCAGATATTCATAGCAACAATATCAGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 261 GluLeuGlyPheHisSerAsnAsnIleArgSerIleProGluLysAlaPheValGlyAsn 280
Qy 841 CCTTCTCTATTACAAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 900
Db 281 ProSerLeuIleThrIleHisPheTyrAspAsnProIleGlnPheValGlyArgSerAla 300
Qy 901 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTGAATGGTGGCTCAACAATACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTCTCTCAAAACCGTCTGCAATCAGTTACCTAAATCTCCAAGTCTAGATCTGTCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACCTATTAGAAGATTACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAATTCAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAACACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTCTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTGGCAAAAATTCATATTATTCACCCCAATGCAATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420

Qy 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCCTGTCGTCTTTTCCCTATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
Qy 1321 GGGTTACATGGTTAACTCACTTAAAAATTAAACAGAAATCATGCTTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAAATCTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAGTGTCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Qy 1441 GCATTTGAGTGTGAGAAATGCTTATAGATTTCTAATCAATGGAATAAGGTGACAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCAGTATGACGACCTTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAATTTCTGCTTGTGACTTTGAGAAAGACCTGAAAGCCCTTCATTCACTGTCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTCACCTTCCCAGGCCCCCTTCAAAACCTGTGAACACCTGCTGTGATGGCTGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATTCGAGTGTGGACCATAGCAGTTCCTGGCACTTACTTGTATGCTTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGTCAATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGGCTGTGTGGATGCTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGGCAGCTTTGCACGACATGCTGCTGGTGGAGAAATGGGGTTGGTGCATGTCAAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTTGCATTTTGTCTGAGATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaLeu 660
Qy 1981 GAGCGTGGTTCCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAAGTAATCATTTTGTCTGTGCCCTGCTGCTGACCATGGCGCAGTTCCTCCCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGGTGGCAGCAGTATGGCGCTCCCTCTCTGCTGCTGCTTTGCTTTTGGGGAGCCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCACCATGGCTACATGGTCTCATCTTGTCTCAATTCCTTGTCTTCTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGCAAGGGAGACCTGGAGAAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Qy 2281 GACTGCTCTATGGTAAACACATTCCTGCTGCTTCTTCCAACTGCATCTCTAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780

QY 901 TTTCAACATTACCTGAACACTAAGACACTGACTCTGAATGCTGCTCACAATAACTCAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAenGlyAlaSerGlnIleThrGlu 320
QY 961 TTTCTGATTAACTGGAACTGCAAACTGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuGlyAlaGlnIle 340
QY 1021 TCATCTCTTCTCAAAACCGTGTGCAATCAGTTACCTAAATCTCCAAGTCTAGATCTGTCT 1080
Db 341 SerSerLeuProGlnThrValCysAenGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
QY 1081 TACAACCTATTAGAGATTACCCAGTTTTCAGTCTGCCAAAGCTTCAGAAATATGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
QY 1141 CTAAGACATAATAGAAATCTAGCAATTAAGTTGACACTTTCCAGCAGTTCCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
QY 1261 TTGCCATCCCTAATAAGCTGACCTATCGTCCACCTCTGCTGCTTTTCTTATTAAT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
QY 1321 GGGTTACATGGTTTAACTCACCTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
QY 1381 TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
QY 1441 GCATTTGAGTGTGTGAGAACTGCTATAGATTTCTAATCAATGAATAAAGGTGACAAAC 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
QY 1501 AGCAGTATGAGCAGCTTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAACGT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
QY 1561 GACCTTGAAGATTTCCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCACTGACG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
QY 1621 TGTTCACTTCCCAAGGCCCTTCAAAACCCCTGTGAACACCTGCTTGAATGGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
QY 1681 AGAATTGGAGTGTGGACCATAGCAGTCTTGGCAGCTTACTTGTAAATGCTTTGGTGAATTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
QY 1741 ACAGTTTTCAGATCCCTCTGATATTTCCTCCCATTAACCTGTTAATTTGGGGTCACTCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
QY 1801 CGAGTGAACATGCTCACGGAGTCTCCAGTGGCTGGCTGGTGGTGGATCGGTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
QY 1861 TTGGCAGCTTTGACGACATGGTCCCTGGTGGGAGAATGGGGTTGGTGGCATGCTATT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
QY 1921 GGTTTTTTGTCCATTTTGTCTCAGAAATCACTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu 660

QY 1981 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCCCCTGTGACCCTTGACCATGGCGGAGTTCCCTGT 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
QY 2101 CTGGGTGGCAGCAGTATGGCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuLeuCysLeuProLeuPheGlyGluPro 720
QY 2161 AGCAACCATGGCTACATGGTGGCTCTCATCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
QY 2221 ACCATTGCTTACACCAAGCTCTACTGCAATTTTGACAAGGGAGACCTGGAGAATATTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
QY 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCCACCAACTGCATCTCTAAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
QY 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
QY 2401 ATTAAGTTTATCTTCTGTTGTTAGTCCCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
QY 2461 ATCTTGTTCATCTCTCACTTTAAGGAGGATCTGTGAGCTGTGAGAGCAAACTACCTGCTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
QY 2521 TGGAACAAGATCAAAACACCAAGCTTGTATGTCAATTAATCTGTATGTGCGAATAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
QY 2581 TCCTGTGACTCAACTCAAGCTTGGTAACTTTACAGCTCCAGCAGTCACCTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyrAspLeu 880
QY 2641 CTTCCCACTTCCGTCGCTCACCACTGCTTATCCAGTGTGAGAGTGCATCTTCTCTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
QY 2701 GTGGCATTTGTCCCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907
RESULT 8
US-10-751-736-84
; Sequence 84, Application US/10751736
; Publication No. US20040265230A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Martinez, Robert
; APPLICANT: Brown, Eugene
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING AND TREATING COLON
; FILE REFERENCE: AM100927 (031896-002000)
; CURRENT APPLICATION NUMBER: US/10/751,736
; PRIORITY FILING DATE: 2003-01-06
; PRIOR APPLICATION NUMBER: US Provisional Application 60/438,000
; PRIORITY FILING DATE: 2003-01-06
; NUMBER OF SEQ ID NOS: 54873
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 84
; LENGTH: 907
; TYPE: PRT

661	Db	GluArgGlyPheSerValTysTySerAlaLysPheGluThrTysAlaProPheSerSer	680
2041	Qy	CTGAAAGTAATCATTTTGCTCTGTGGCCCTGTCTGGCCCTTGACCATGGCGCGAGTGTCCCTCG	2100
681	Db	LeuLysValIleIleLeuLeuCyAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu	700
2101	Qy	CTGGGTGGCAGCAAGTAGTATGGCGCTCCCTCTCTGCGCTGCGCTTTGGCTTTGGGGAGCC	2160
701	Db	LeuGlyGlySerLysTyGlyAlaSerProLeuCySLeuProLeuProPheGlyGluPro	720
2161	Qy	AGCACCATGGCTACATGGTGGCTCTCATCTTGCTCAATCCCTTTGCTTCCTCATCATG	2220
721	Db	SerThrMetGlyTyMetValAlaLeuIleLeuLeuAsnSerLeuCySPheLeuMetMet	740
2221	Qy	ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTCGGAGAATATTTGG	2280
741	Db	ThrIleAlaTyThrLysLeuTyCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr	760
2281	Qy	GACTGTCTATGGTAAACACATTTGCCCTGTGTGCTTCCACCACTGCATCCTTAACTGC	2340
761	Db	AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys	780
2341	Qy	CCTGTGGCTTTCTGTGCTCTCTCTCTTTATAAACCTTACATTTATCAGTCTCGTGAAGTA	2400
781	Db	ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal	800
2401	Qy	ATTAAGTTTATCTTCTGGTGGTAGTCCCACCTCTCTGATGTGTCAATTCGCTTCTCTAC	2460
801	Db	IleLysPheIleLeuLeuValValValProLeuProAlaCysLeuAsnProLeuLeuTyr	820
2461	Qy	ATCTTGTTTCATCTCATTTTAAAGGAGATCTGGTGAGCCTGAGAAAGCAACCTACGTC	2520
821	Db	IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal	840
2521	Qy	TGGACAAGATCAAAACACCCAGCTTGATGTGCAATTAACTCTGATGATGTGCAAAAACAG	2580
841	Db	TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln	860
2581	Qy	TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTACCAGCTCCACATCACTTATGACCTG	2640
861	Db	SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu	880
2641	Qy	CCTCCCAGTTCCGTGGCCATCACCAAGCTTATCCAGTGTAGAGAGCTGCCATCTTCTCTCT	2700
881	Db	ProProSerSerValProSerProAlaTyProValThrGluSerCysHisLeuSerSer	900
2701	Qy	GTGGCATTTGTCCCAGTCTC	2721
901	Db	ValAlaPheValProCysLeu	907

```

RESULT 9
US-10-482-029-158
; Sequence 158, Application US/10482029
; Publication No. US20050037445A1
; GENERAL INFORMATION:
; APPLICANT: ODIN medical A/S
; TITLE OF INVENTION: Oncology drug innovation
; FILE REFERENCE: P 573 PC00
; CURRENT APPLICATION NUMBER: US/10/482.029
; CURRENT FILING DATE: 2003-12-29
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 158
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-482-029-158

Alignment Scores:
Pred. No.: 0
Score: 907.00
Percent Similarity: 100.00%
Conservative: 0
Length: 907
Matches: 907

```

Qy	961	TTTTCTCATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC	1020
Db	321	PheProAspLeuThrGlyIvThrAlaAsnLeuGluSerLeuThrGlyAlaGlnIle	340

1021 TCATCTCTTCTCAAAACCGTCTGCAATCAAGTACCTAAATCTCAAGTGTCTAGATCTGCT 1080
Db SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
1081 TACAACCTATTAGAGATTACCCAGTCTTTCAGTCTGCGCAAAAGCTTCAGAAAATTGAC 1140
Db TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLeuGlnLysIleAsp 380
1141 CTAAGACATAATCAAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTCTTAGCCTC 1200
Db LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
1201 CGATCGCTCAATTTGGCTTGGAAACAAAATGCTATTATTATTCACCCCAATGCAATTTCCACT 1260
Db ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
1261 TTGGCATCCCTAATAAGCTGACACTATCGTCCAACTCTCTGCTCTTCTCTTACT 1320
Db LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
1321 GGGTTACATGGTTAACTCACTTAAATTAACAGGAAATCATGCGCTTACAGAGCTTGATA 1380
Db GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
1381 TCATCTGAAAACCTTTCCAGAACTCAAGTTATAGAAATGCCCTTATGCTTACCAGTCTGT 1440
Db SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCys 480
1441 GCATTTGGAGTGTGAGAGATCCCTATAGATTTCTAATCAATGAATAAAGTGCACAC 1500
Db AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
1501 AGCAGTATGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAAGCT 1560
Db SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
1561 GACCTTGAAAGATTTCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
1621 TGTTTCACTTCCCGAGGCCCTTCAACCCCTGTGAACACTCTGTGATGGCTGGCTGATC 1680
Db CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
1681 AGAATTTGGAGTGTGACCATAGCAGTCTGGCACTTACTTGTAAATGCTTGTGACTTCA 1740
Db ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTAATGAGGCTCATCGCA 1800
Db ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
1801 GCAGTGAACATGCTCAGGGAGTCTCAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGACT 1860
Db AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
1861 TTTGGCAGCTTGGCAGCATGCTGCTGGTGGGAGATGGGGTGGTGGCTGCTGCTGCTGCT 1920
Db PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
1921 GGTTTTTGTCCATTTTCTTCCAGATCATCTGTTTTCTGCTTCTTCTTCTTCTGTCAGCCCTG 1980
Db GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
1981 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
Db GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
2041 CTGAAAGTAAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700

2101 CTGGTGGCAGCAGATATGGGCGCTCCCTCTCTGCTGCTGCTTTCGCTTTGGGAGCCCC 2160
Db LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
2161 AGCACCATTGGCTACATGCTGCTCTCATCTTGTCAATTCCTTTGCTTCTCTCATGATG 2220
Db SerThrMetGlyTyrMetValAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuMetMet 740
2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGATAATTTGG 2280
Db ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTyr 760
2281 GACTGCTCTATGGTAAAAACACATTCCTGCTGCTTCCACCACTGCATCTTAACCTGC 2340
Db AspCysSerMetValLysHisIleAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuAsnCys 780
2341 CCTGGGCTTCTCTGCT 2400
Db ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
2401 ATTAAGTTTATCT 2460
Db IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
2461 ATCTGTTCAATCTCTCACTTTTAAGGAGATCTGTGAGCTGAGAAAGCAAACTACGTC 2520
Db IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
2521 TGGCAACATCAAAACACCAAGCTTGTGATGCAATTAATCTGTGATGATGTCGAAAAACAG 2580
Db TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspValGluLysGln 860
2581 TCCTGTGACTCAACTCAAGCTTGTGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
Db SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
2641 CTCCAGCTTGGTGGCATCCAGCTTATCCAGTGCAGTGCAGCTGCCATCTTCTCTCT 2700
Db ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
2701 GTGGCATTTGCTCCCATGCTC 2721
Db ValAlaPheValProCysLeu 907

RESULT 10
US-10-251-385-278
; Sequence 278, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251.385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170.496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 278
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-251-385-278

Alignment Scores:
Pred. No.: 0 Length: 907
Score: 806.00 Matches: 906
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1

Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACTATTAGAGATTACCCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAAGACATATGAATCTACGAATTAAGCTTACACTTTCCAGCAGTCTCTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCGTATTATTACCCCAATGCAATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleHisProAsnAlaPheSerThr 420
Qy 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCTCGTCTTTTCCCTATTAAC 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThr 440
Qy 1321 GGGTTACATGGTTTAACCTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGTGTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys 480
Qy 1441 GCATTTGGAGTGTGAGAAATGCCCTAAGATTTCTAATCAATGGAATAAAGGTGACAA 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCAGTATGGAGCCTTCATAAGAAAGATCTGGAATGTTTCAGGCTCAAGATGAAGCT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAGATTTCTCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCATTTCAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTCACCTTCCCGAGGCCCTTCAACCCCTGTGAAACACCTGCTTGATGGCTGTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATTGAGTGTGGACCATAGCAGTCTGGCAGCTTACTGTGAATGCTTGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 GCAGTGAACATGCTCACGGAGTCTCCAGTCCGCTGGCTGGCTGGTGGATCGCTTCACT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTTGACAGCTTGCAGCAGATGCTCCCTGTTGGGAGAAATGGGTTGGTCCATGCTCAT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGCGACCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu 660
Qy 1981 GAGCTGGGTCTCTGTGGAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAGTAATCATTTGCTCTGCTCCCTGCTGGCTTGGACCATGCGCGAGTTCCTCGT 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGGTGGCAGCAGATGAGCGCCCTCCCTCTCTGCTGCTTTCCTTTTGGGAGGCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuPheGlyGluPro 720

Qy 2161 AGCACCATGGGTACATGGTGGCTCTCATCTTGCTCAATTCCTTCTTCTCTCATGAT 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTCCCTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Qy 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGTCTTCCAACTGATCCTTAAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Qy 2341 CCTGTGGCTTCTGTCTCTCTCTCTTAACTTAACCTTACATTTATCAGTCCCTGGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Qy 2401 ATTAAGTTTATCCCTTCTGGTGGTAGTCCCACTTCTCTGCATGTCTCAATCCCTTCTCTAC 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Qy 2461 ATCTTGTTCATCTCTCTTTAAGGAGATCTGTGAGCTGTGAGCTGTGAGAAAGCAACCTACGTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyrVal 840
Qy 2521 TGGACAGATCAAAACACCCAGCTTCTGATGTCAATTAACCTCTGATGATGTCGAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Qy 2581 TCCTGTGACTCAACTCAAGCCTTGTAACTTTTACCAGCTCCAGCATCATTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyrAspLeu 880
Qy 2641 CCTCCCAGTTCCTGTGCCATCACAGCTTATCCAGTTCAGTGTGAGCTGAGCTGCTTCTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Qy 2701 GTGGCATTTGTCCCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 12
US-09-970-944-27
; Sequence 27, Application US/0970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Herrman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shimkets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; FILE OF INVENTION: Antibodies Directed Against these Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 27
; LENGTH: 907
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-27

Alignment Scores:
Pred. No.: 0 Length: 907
Score: 716.00 Matches: 816
Percent Similarity: 99.76% Conservative: 0
Best Local Similarity: 99.76% Mismatches: 1
Query Match: 78.94% Indels: 2
DB: 10 Gaps: 0
US-10-751-736-21 (1-2724) x US-09-970-944-27 (1-907)

QY 271 TTCCTGGAGAGTACGCTCTTGGGGAACGCTCTGACATACATATCCCAAGGAGCATTC 330
 Db |||||
 QY 91 PheLeuGluGluLeuArgLeuAlaGlyAsnAlaLeuThrTyrIleProLysGlyAlaPhe 110
 Db |||||
 QY 331 ACTGGCCTTTACAGCTCTTAAAGTTCTTATGCTGCAGATAAATCAGCTAAGACACGCTACCC 390
 Db |||||
 QY 111 ThrGlyLeuTyrSerLeuLysValLeuMetLeuGlnAsnGlnLeuArgHisValPro 130
 Db |||||
 QY 391 ACAGAAGCTCTCAGAAATTTGCAAGCCTTCAATCCCTGCTGCTGATGCTAACCACATC 450
 Db |||||
 QY 451 ACCTATGCCCCCAAGCTGTTTCAGTGCCCTGCATTCCTCAGGACCTGTGGCTGGAT 510
 Db |||||
 QY 151 SerTyrValProProSerCysPheSerGlyLeuHisSerLeuArgHisLeuTrpLeuAsp 170
 Db |||||
 QY 511 GCAATGCGCTTAACAGAAATCCCGTCCAGGCTTTTAGAAGTTTATCGSCATTTCGAAGCC 570
 Db |||||
 QY 171 AspAsnAlaLeuThrGluIleProValGlnAlaPheArgSerLeuSerAlaLeuGlnAla 190
 Db |||||
 QY 571 ATGACCTTGGCCCTGAACAAAATACACACATACCAGACTATGCTCTTGGAAACCTCTCC 630
 Db |||||
 QY 191 MetThrLeuAlaLeuAsnLysIleHisIleProAspTyrAlaPheGlyAsnLeuSer 210
 Db |||||
 QY 631 AGCTTGSTA-GTTCTACATCTCCATAA CAATAGAATCCACTCCCTGGGAAAGAAATGCTT 689
 Db |||||
 QY 211 Ser-TrpValValLeuHisLeuHisAsnAsnArgIleHisSerLeuGlyLysLysCysPh 230
 Db |||||
 QY 690 TGATGGCTCCACACCTAGACACTTATAGATTAAATACATAATACCTTGATGAATCCC 749
 Db |||||
 QY 230 eAspGlyLeuHisSerLeuGluThrLeuAspLeuAsnTyrAsnAsnLeuAspGluPhePr 250
 Db |||||
 QY 750 CACTGCAATTAGGACACTCTCAACCTTAAAGAACTAGGATTTTCATAGCAACAATATCAG 809
 Db |||||
 QY 250 ofhAlaIleArgThrLeuSerAsnLeuLysGluLeuGlyPheHisSerAsnAsnIleAr 270
 Db |||||
 QY 810 GTCGATACCTGAGAAAGCATTTGTAGGCAACCCCTCTCTTATTA CAATACATTTCTATGA 869
 Db |||||
 QY 270 gSerIleProGluLysAlaPheValGlyAsnProSerLeuIleThrIleHisPheTyrAs 290
 Db |||||
 QY 870 CAATCCCATCAATTTGTTGGAGATCTGCTTTTCAACATTTACCTGAACTAAGAACACT 929
 Db |||||
 QY 290 pAsnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGluLeuArgThrLe 310
 Db |||||
 QY 930 GACTCTGAATGTGCTCACAAATAACTCAATTTCTGATTTAACTGGAACCTGCAACCT 989
 Db |||||
 QY 310 uThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGlyThrAlaAsnLe 330
 Db |||||
 QY 990 GGAGAGCTGACTTTAACTGGAGCA CAGATCTCATCTCTTCTCAAAACGCTCTGCAATCA 1049
 Db |||||
 QY 330 uGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThrValCysAsnG1 350
 Db |||||
 QY 1050 GTTACTATCTCCAAGTGTAGTCTGCTTACAACTTATAGAGATTACCCAGTTT 1109
 Db |||||
 QY 350 nLeuProAsnLeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAspLeuProSerPh 370
 Db |||||
 QY 1110 TTCAGTCTGCCAAAAGCTTCAGAAAATTCAGCAATATGAAATCTACGAAATTA 1169
 Db |||||
 QY 370 eSerValCysGlnLysLeuGlnLysIleAspLeuArgHisGlnGluIleTyrGluIleLys 390
 Db |||||
 QY 1170 AGTTGACACTTTCCAGCAGTGTCTTAGCTCCGATCCGATCGCTGGAATTTGGCTTGGAAACAAAT 1229
 Db |||||
 QY 390 sValAspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAlaTrpAsnLysI1 410
 Db |||||
 QY 1230 TGCTATTATTCACCCCAATGCAATTTTCCATTTGGCATCCCTTAATAAGCTGGACCTATC 1289
 Db |||||
 QY 410 eAlaIleIleHisProAsnAlaPheSerThrLeuProSerLeuIleLysLeuAspLeuSe 430
 Db |||||
 QY 1290 GTCCAACTCTCTGCTGCTTTTCTTATAACTGGGTTTACATGGTTTAACTCACTTAAATTT 1349
 Db |||||
 QY 430 rSerAsnLeuLeuSerSerPheProIleThrGlyLeuHisGlyLeuThrHisLeuLysLe 450
 Db |||||

QY 1350 AACAGAAATCATGCTTACAGAGCTTGATATCATCTGAAACTTTTCGAAACTCAAGGT 1409
 Db |||||
 QY 450 uThrGlyAsnHisAlaLeuGlnSerLeuIleSerGluAsnPheProGluLeuLysVa 470
 Db |||||
 QY 1410 TATCAAAATGCCCTTATGCTTACCAGTGTCTGTCATTTGGAGTGTGTGAGAATGCCCTATAA 1469
 Db |||||
 QY 470 I1leGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGluAsnAlaTyrLys 490
 Db |||||
 QY 1470 GATTTCTTAATCAATGGAATAAAGGTGCAACACAGCATGTGGACGACCTTTTCATAAGAAAGA 1529
 Db |||||
 QY 490 sIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspLeuHisLysLysAs 510
 Db |||||
 QY 1530 TGCTGGAATGTTTCAAGCTCAAGATGAACGTCACCTTGAAGATTTCTGCTTGACTTTGA 1589
 Db |||||
 QY 510 pAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeuLeuAspPheG1 530
 Db |||||
 QY 1590 GGAACACCTGAAAGCCCTTTCATTCAGTCAGTGTTCACCTCCCGAGGCCCTTCAAAACC 1649
 Db |||||
 QY 530 uGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGlyProPheLysPr 550
 Db |||||
 QY 1650 CTGTGAACACCTGCTGTGATGGCTGATCAGAAATTGGAGTGTGGACCATAGCAGTTCT 1709
 Db |||||
 QY 550 oCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThrIleAlaValLe 570
 Db |||||
 QY 1710 GGCACCTACTTGTAAATGCTTTGGTGACTTCAACAGCTTTTTCAGATCCCTCTGTACATTC 1769
 Db |||||
 QY 570 uAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerProLeuTyrIleSe 590
 Db |||||
 QY 1770 CCCCAATTAACCTGTTAATTTGGGCTCATCGACAGCTGAACATGCTCAGGGAGTCTCCAG 1829
 Db |||||
 QY 590 rProIleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThrGlyValSerSe 610
 Db |||||
 QY 1830 TGCCGTGCTGCTGTGTGATGCTTTCACATTTTGGAGCTTTGACGACCATGGTGCCTG 1889
 Db |||||
 QY 610 rAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArgHisGlyAlaTr 630
 Db |||||
 QY 1890 GTGGAGAAATGGGTTGTTGCCATGTCATTTGGTGTGTTTTTGGCCATTTTGTCTTCAATC 1949
 Db |||||
 QY 630 pTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePheAlaSerGluSe 650
 Db |||||
 QY 1950 ATCTGTTTTCTGCTTACTCTGGCAGCCCTGGAGGCTGGTTCCTGTGTAATATTTCTGC 2009
 Db |||||
 QY 650 rSerValPheLeuLeuThrLeuAlaLeuGluArgGlyPheSerValLysTyrSerAl 670
 Db |||||
 QY 2010 AAAATTTGAAACGAAAGCTCCATTTTTCAGCTCTGAAAGTAAATCATTTTGTCTGTGCCCT 2069
 Db |||||
 QY 670 aLysPheGluThrLysAlaProPheSerSerLeuLysValIleLeuLeuCysAlaLe 690
 Db |||||
 QY 2070 GCTGGCTTGACCATGGCCGAGTTCCCTCTGCTGGTGCGCAAGATATGGCGCTCCCC 2129
 Db |||||
 QY 690 uLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyrGlyAlaSerPr 710
 Db |||||
 QY 2130 TCTCTGCTGCTGCTTTCCTTTTGGGAGCCAGACCATGGCTACATGGTGCCTCAT 2189
 Db |||||
 QY 710 oLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMetValAlaLeuI1 730
 Db |||||
 QY 2190 CTGTGCTCAATTTCCCTTTTCTCTCATGATGACCATTCCTACACCAAGCTCTACTGCAA 2249
 Db |||||
 QY 730 eLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLysLeuTyrCysAs 750
 Db |||||
 QY 2250 TTTGGAACAGGAGACCTGGAGAAATATTTGGGAGCTGCTGTATGGTAAACACATGGCCCT 2309
 Db |||||
 QY 750 nLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLysHisIleAlaLe 770
 Db |||||
 QY 2310 GTTGCTCTTCAACCACTGCATCCATAACCTGCCCTGTGGCTTCTTGTCTCTCTCTCTTT 2369
 Db |||||
 QY 770 uLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSerPheSerSerLe 790
 Db |||||
 QY 2370 AATAAACCTTACATTTTACCTCTGAAAGTAAATTAAGTTTATCTCTTCTGGTGGTAGTCCC 2429
 Db |||||
 QY 790 uIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeuValValPr 810
 Db |||||
 QY 2430 ACTTCTGTCATGCTCAATCCCTTCTCTCATCATCTTGTTCATCTCTCACTTTAAGGAGGA 2489
 Db |||||

Db 810 oLeuProAlaCysLeuAsnProLeuLeuTyRileLeuPheAsnProHisPheLysGluAs 830
Qy 2490 TCTGGTGGCCCTGAGAAACCAACCTACGCTGGACCAAGATCAAAACACCCCAAGCTTGCAT 2549
Db 830 pLeuValSerLeuArgLysGlnThrTyRValTrpThrArgSerLysHisProSerLeuMe 850
Qy 2550 GTCATTAACCTCTGATGATGTCGAAAAACAGTCCTGTGACCTCAACTCAGCCTTGGTAAC 2609
Db 850 tSerIleAsnSerAspValGluLysGlnSerCysAspSerThrGlnAlaLeuValTh 870
Qy 2610 CTTTACCAAGCTCCAGCATCTTATGACCTGCTCCCAAGTTCCGTCGACATCACCAGCTTA 2669
Db 870 rPheThrSerSerSerIleThrTyRAspLeuProProSerSerValProSerProAlaTy 890
Qy 2670 TCCAGTGACTGAGAGCTGCCATCTTCTCTGTGGCAATTTGTCACATGCTTC 2721
Db 890 rProValThrGluSerCysHisLeuSerSerValAlaPheValProCysLeu 907

RESULT 13

US-10-295-027-484
; Sequence 484, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 484
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-484

Alignment Scores:
Pred. No.: 0 Length: 883
Score: 621.00 Matches: 621
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 68.47% Indels: 0
DB: 15 Gaps: 0
US-10-751-736-21 (1-2724) x US-10-295-027-484 (1-883)
Qy 859 CATTTCTATGACAAATCCCAATCTGTTGGGAGATCTGCTTTTCAACATTTACCTGAA 918
Db 263 HisPheTyRAspAsnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu 282
Qy 919 CTAGAAACACTGACTCTGAAATGGTCTCAAAATAACTGAATTTCTGTATTTAACTGGGA 978
Db 283 LeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGly 302
Qy 979 ACTCAAAACCTGGAGACTCTGACTTTAACTGGAGCAGATCTCATCTCTCTCCCAACC 1038
Db 303 ThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThr 322
Qy 1039 GTCTGCAATCAGTTTACCTAATCTCCAAGTGTAGATCTGTCTTACAACTTATTAGAAGAT 1098
Db 323 ValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSerTyRAsnLeuLeuGluAsp 342
Qy 1099 TTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGACCTTAAGACATATAAGAAATC 1158
Db 343 LeuProSerPheSerValCysGlnLysLeuGlnLysIleAspLeuArgHisAsnGluIle 362
Qy 1159 TACGAAATTAAGCTTGACACTTCCAGCAGTTGCTTAGCCTCCGATCGCTGAATTTGGCT 1218
Db 363 TyrGluIleLysValAspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAla 382
Qy 1219 TGGAAACAAAATTGCTATTATTATTCACCCCAATGCATTTTCCACTTTGGCCATCCCTAATAAG 1278
Db 383 TrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThrLeuProSerLeuIleLys 402
Qy 1279 CTGGACCTATCGTCCAACTCTCTGCTCTTTCTTATAACTGGGTACATGGTTTAACT 1338
Db 403 LeuAspLeuSerSerAsnLeuLeuSerSerPheProIleThrGlyLeuHisGlyLeuThr 422
Qy 1339 CACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATATCATCTCGAAAACCTTTCCA 1398
Db 423 HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPhePro 442
Qy 1399 GAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAGTGTGTGTGATTTGGAGTGTGAG 1458
Db 443 GluLeuLysValIleGluMetProTyRAlaTyRLeuGlnCysCysAlaPheGlyValCysGlu 462
Qy 1459 AATCCCTATAAGATTTCTAATCGAATAAGGTGACAAACAGCAGTATGGACAGCTT 1518
Db 463 AsnAlaTyRlyLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu 482
Qy 1519 CATAGAAGAGATGCTGGAATGTTTTCAGGCTCAAGATGAAGTGAACCTTTGAAGATTTCTGT 1578
Db 483 HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeu 502
Qy 1579 CTTGACTTTTGGAGAGACCTGAAAGCCCTTCATTTCAGTGCAGTGTTCACCTTCCCCAGGC 1638
Db 503 LeuAspPheGluGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGly 522
Qy 1639 CCCTTCAAAACCTGTGAACACCTGCTGTAGTGGCTGTGATCAGAATTTGGAGTGTGAGC 1698
Db 523 ProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThr 542
Qy 1699 ATAGCAGTTCTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCAACAGTTTTTCAGATCCCT 1758
Db 543 IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro 562
Qy 1759 CTGTACATTTTCCCCCATTAACCTGTTAATGGGTCTATCGCAGCAGTGAACATGCTCACG 1818
Db 563 LeuTyRileSerProIleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr 582
Qy 1819 GGAGTCTCCAGTCCGCTGCTGGTGTGATGCGTTTCACTTTTGGCAGCTTTGACGAG 1878
Db 583 GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg 602

1279 CTGGACCTATCGTCCAACTCTGTCGCTTTTCTCTATAAAGTGGTTACATGTTTAACT 1338
Db |||||||
403 LeuAspLeuSerSerAsnLeuLeuSerPheProIleThrGlyLeuHisGlyLeuThr 422
Qy |||||||
1339 CACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATATCATCTGAAAACTTTTCCA 1398
Db |||||||
423 HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPhePro 442
Qy |||||||
1399 GAATCAAGGTTATAGAAATGCCCTTATGCTTACCAAGTCTGTGCAATTTGGAGTGTGAG 1458
Db |||||||
443 GluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGlu 462
Qy |||||||
1459 AATGCCATATAGATTTCTAATCAATGGAATAAGGTGACACAGCAGTATGGACGACCTT 1518
Db |||||||
463 AsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspLeu 482
Qy |||||||
1519 CATAGAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGTGACTTGAAGATTTCTCTG 1578
Db |||||||
483 HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeu 502
Qy |||||||
1579 CTTGACTTTGAGGAGACCTGAAAGCCCTTCAATCAGTGCAGTGTTCACCTTCCCCAGGC 1638
Db |||||||
503 LeuAspPheGluGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProGly 522
Qy |||||||
1639 CCCTTCAACCCCTGTGAACACCTGCTGATGGCTGGCTGATCAGAAATGGAGTGGACC 1698
Db |||||||
523 ProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThr 542
Qy |||||||
1699 ATAGCAGTTCGGCACTTACTTGTAAATGCTTTGGTGACTTCAACAGTTTTTCAGATCCCT 1758
Db |||||||
543 IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro 562
Qy |||||||
1759 CTGTACATTTCCCCATTTAAACTGTTAATGGGTTCATCGCAGCAGTGAACATGCTCAG 1818
Db |||||||
563 LeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr 582
Qy |||||||
1819 CGAGTCTCAGTGCCTGCTGGCTGGTGTGATGCTGCTTCACTTTTGGCAGCTTTGGCAGA 1878
Db |||||||
583 GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg 602
Qy |||||||
1879 CATGTCCTGCTGGGAGAATGGGTTCGTTGCCATGTCAATGGTTTTCATCATTTT 1938
Db |||||||
603 HisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePhe 622
Qy |||||||
1939 GCTTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTGGAGCGGTGTTCTCTGTG 1998
Db |||||||
623 AlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerVal 642
Qy |||||||
1999 AATATTTCTGCAAAATTTGAAACGAAGCTCCATTTCTAGCCTGAAAGTAAATCATTTTG 2058
Db |||||||
643 LysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValIleLeuLeu 662
Qy |||||||
2059 CTCTGTGCCCTGCTGGCTTGACCATGGCCGAGTTCCTGCTGGGTGGGAGCAAGTAT 2118
Db |||||||
663 LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyr 682
Qy |||||||
2119 GGCGCTCCCTCTCTGCTGCTGCTTGTGCTTTTGGGAGCCGACCATGGCTTACATG 2178
Db |||||||
683 GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet 702
Qy |||||||
2179 GTCGCTCTCATCTTCTCAATCCCTTTGCTTCTCATGATGACCATTTGCTTACACCAAG 2238
Db |||||||
703 ValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLys 722
Qy |||||||
2239 CTCTACTGCAATTTGGCAAGGAGACCTGAGAGATATTTGGGACTGCTCTATGTTAAA 2298
Db |||||||
723 LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLys 742
Qy |||||||
2299 CACATGCGCTGTGCTCTTCAACCAACTGCAATCTGCAATCTGCTGCTGCTGCTGCTGCT 2358
Db |||||||
743 HisIleAlaLeuLeuLeuPheThrAsnCysLysIleLeuAsnCysProValAlaPheLeuSer 762

Qy |||||||
2359 TTCTCCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTAAATTAAGTTTATCTCTG 2418
Db |||||||
763 PheSerSerLeuLeuAsnLeuThrPheIleSerProGluValIleLysPheIleLeuLeu 782
Qy |||||||
2419 GTGCTAGTCCACCTTCTCGCATGTCTCAATCCCTTCTTACATCTTTGTTCAATCTCTCAC 2478
Db |||||||
783 ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis 802
Qy |||||||
2479 TTTAAGAGAGATCTGGTGAGCCTGAGAAAGCAACCTACGCTGAGCAAGATCAAAACAC 2538
Db |||||||
803 PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTrpThrArgSerLysHis 822
Qy |||||||
2539 CCAAGCTTGATGTCAATTAATCTCTGATGATGTCGAAAAACAGTCTCTGTGACTCAACTCAA 2598
Db |||||||
823 ProSerLeuMetSerIleAsnSerAspValGluLysGlnSerCysAspSerThrGln 842
Qy |||||||
2599 GCCTTGGTAACTTTTACCAGTCCAGCATCACTTATGACCTGCTCCAGTTCCTGCTGCTG 2658
Db |||||||
843 AlaLeuValThrPheThrSerSerSerIleThrTyrAspLeuProProSerSerValPro 862
Qy |||||||
2659 TCACGACTTATCCAGTGCAGTGCAGAGTGCATCTTCTCTGCTGCGCATTTGTTCCCATGT 2718
Db |||||||
863 SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys 882
Qy |||||||
2719 CTC 2721
Db |||||
883 Leu 883
RESULT 15
US-10-173-999-28
; Sequence 28, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10173,999
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 883
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-28
Alignment Scores:
Pred. No.: 0 Length: 883
Score: 621.00 Matches: 621
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 68.47% Indels: 0
DB: 15 Gaps: 0
US-10-751-736-21 (1-2724) x US-10-173-999-28 (1-883)
Qy 859 CATTTCTATGATCCATCCCATTTTGTGGGAGATCTGCTTTTCAACATTTACTTGAA 918
Db |||||||
263 HisPheTyrAspAsnProIleGlnPheValGlyArgSerAlaPheGlnHisLeuProGlu 282
Qy |||||||
919 CTAAGAACACTGACTCTGAATGGTGCCTCAACAATAAATGAATTTTCTGATTTAACTGGA 978

|||||
283 LeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGluPheProAspLeuThrGly 302
QY ACTGCAAACTGGAGAGTCTGACTTAACTGGAGCACAGATCTCATCTCTCTCAAAACC 1038
Db |||||||
303 ThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIleSerSerLeuProGlnThr 322
QY GTCTGCAATCAGTACTTAATCTCAAGTGTAGATCTGTCTTTACAACTATTAGAGAT 1098
Db |||||||
323 ValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSerTyrAsnLeuLeuGluAsp 342
QY TTACCCAGTTTTTCACTGCGCAAAAGCTTCGAAATTCAGCTAAGACATAGAAATC 1158
Db |||||||
343 LeuProSerPheSerValCysGlnIysLeuGlnIysLeuGlnIysLeuAspLeuArgHisAsnGluIle 362
QY TAGCAAAATTAAGTTCACACTTTCCAGCAGTTCCTTAGCTCCGATCGCTGAATTTGGCT 1218
Db |||||||
363 TyrGluIleValAspThrPheGlnGlnLeuLeuSerLeuArgSerLeuAsnLeuAla 382
QY TGAACAAAATTCGTATTATTCACCCCAATGCAATTTCCATTTGCCATCCCTAATAAG 1278
Db |||||||
383 TrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThrLeuProSerLeuIleIys 402
QY CTGGACCTATGTCCAACCTCTGCTGCTTTTCTTAATCTGAGTGGGTACATGGTTAACT 1338
Db |||||||
403 LeuAspLeuSerSerAsnLeuLeuSerPheProIleThrGlyLeuHisGlyLeuThr 422
QY CACTTAAATTAACAGAAATCATGCTTACAGAGCTTCATATCATCTCAAACTTTCCA 1398
Db |||||||
423 HisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIleSerSerGluAsnPhePro 442
QY GAACCAAGGTTATAGAAATGCTTATGCTTACCAGTGTGTCATTTGGAGTGTGAG 1458
Db |||||||
443 GluLeuLysValIleGluMetProTyrAlaTyrGlnCysCysAlaPheGlyValCysGlu 462
QY ATGCTTATAAGATTTCTAATCAATGGAATAAGGTGCAACAGCAGTATGACGACCTT 1518
Db |||||||
463 AsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsnSerSerMetAspAspLeu 482
QY CATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAGTCAAGCTTCAAGATTTCTGT 1578
Db |||||||
483 HisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArgAspLeuGluAspPheLeu 502
QY CTGACTTTGAGAAAGACCTGAAAGCCCTTCATTCAGTGCAGTGTTCACCTTCCCAGGC 1638
Db |||||||
503 LeuAspPheGluGluAspLeuLysAlaLeuHisSerValGlnCysSerProSerProgly 522
QY CCTTTCMAACCTGTGGAACACTGTTGATGCTGGCTGCATCAGAAATTCGAGTGTGACC 1698
Db |||||||
523 ProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIleArgIleGlyValTrpThr 542
QY ATAGCAGTTCGGCACTTACTTGTATGCTTTGTTGGTGAATTCACAGTTCAGATCCCT 1758
Db |||||||
543 IleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSerThrValPheArgSerPro 562
QY CTGTACATTTCCCCATTAATGTTTAAATTTGGGGTCATCCGACAGTGAACATGCTCAG 1818
Db |||||||
563 LeuTyrIleSerProIleLysLeuLeuIleGlyValIleAlaAlaValAsnMetLeuThr 582
QY GGAGTCTCCAGTCCGCTGCTGGCTGGTGGATGGTTCACCTTTTGGCAGCTTTGCACGA 1878
Db |||||||
583 GlyValSerSerAlaValLeuAlaGlyValAspAlaPheThrPheGlySerPheAlaArg 602
QY CATGGTCTGGTGGGAGAAATGGGGTGTGCTCATGTCATTTGTTTGTTCCTTTT 1938
Db |||||||
603 HisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIleGlyPheLeuSerIlePhe 622
QY GCTTCAGAAATCATCTGTTTCTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTCTCTGTG 1998
Db |||||||
623 AlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeuGluArgGlyPheSerVal 642
QY AAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCTGAAAGTAAATCATTTTG 2058
|||||

Db 643 LysTyrSerAlaLysPheGluThrLysAlaProPheSerSerLeuLysValIleIleLeu 662
QY CTCTGTGGCCTGCTGGCCTTGACCATGGCCGAGTTCCCTGCTGCTGGGTGGCAGCAAGTAT 2118
Db |||||||
663 LeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeuLeuGlyGlySerLysTyr 682
QY GCGGCTCCCTCTCTGCTGCTGCTTTGGGGAGCCCGAGCAACCATGGGCTACATG 2178
Db |||||||
683 GlyAlaSerProLeuCysLeuProLeuProPheGlyGluProSerThrMetGlyTyrMet 702
QY GTGCGCTCATCTGCTCAATTCCTTTGCTTCTCATGATGACCATTCCTTACCAACG 2238
Db |||||||
703 ValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMetThrIleAlaTyrThrLys 722
QY CTCTACTGCAATTTGGACAAGGAGACCTGGGAAATATTTTGGGACTGCTCTTATGGTAA 2298
Db |||||||
723 LeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrpAspCysSerMetValLys 742
QY GACATTTGCCCTGCTGCTTTCACCAATGTCATCTTAAACTGCCCTGTGGCTTTCTTGTC 2358
Db |||||||
743 HisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCysProValAlaPheLeuSer 762
QY TTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTGAGTAATTAAGTTTATCCTTCTG 2418
Db |||||||
763 PheSerSerLeuIleAsnLeuThrPheIleSerProGluValIleLysPheIleLeu 782
QY GTGGTAGTCCCACTTTCCTGCTGCTCAATCCCTTCTCTACATCTTGTTCATCTCCAC 2478
Db |||||||
783 ValValValProLeuProAlaCysLeuAsnProLeuLeuTyrIleLeuPheAsnProHis 802
QY TTTAAGGAGGATCTGCTGAGCCTGAGAAAGCAAACTAGCTCTGGACAAGATCAAAACAC 2538
Db |||||||
803 PheLysGluAspLeuValSerLeuArgLysGlnThrTyrValTrpThrArgSerLysHis 822
QY CCAAGCTTGATGTCATTAATCACTCATGATGTGAAABACACTCCTGTGACTCAACTCAA 2598
Db |||||||
823 ProSerLeuMetSerIleAsnSerAspAspValGluLysGlnSerCysAspSerThrGln 842
QY GCCTTTGGTAAACCTTTTACCAGCTCCAGCATCACTTATGACCTGCCTCCAGTTCCGTC 2658
Db |||||||
843 AlaLeuValThrPheThrSerSerSerIleThrTyrAspLeuProProSerSerValPro 862
QY TCACAGCTTATCCAGTGCAGTGCAGAGCTGCCATCTTTCTCTGTGGCATTTGTCCCATGT 2718
Db |||||||
863 SerProAlaTyrProValThrGluSerCysHisLeuSerSerValAlaPheValProCys 882
QY CTC 2721
Db |||
883 Leu 883

Search completed: July 13, 2005, 04:08:29

Job time : 331.5 secs

This page Blank (uspto)

SUMMARIES									
Result No.	Query #			DB	ID	Description			
	Score	Match	Length						
1	2724	100.0	2724	6	CQ721366	CQ721366 Sequence			
2	2724	100.0	2724	6	CQ766844	CQ766844 Sequence			
3	2724	100.0	2724	6	AK308745	AK308745 Sequence			
4	2724	100.0	2724	6	AX549136	AX549136 Sequence			
5	2724	100.0	2880	6	AX658241	AX658241 Sequence			
6	2724	100.0	2880	9	AF062006	AF062006 Homo sapi			
7	2724	100.0	4559	6	BD075815	BD075815 G protein			
8	2724	100.0	4570	6	AB447625	AB447625 Sequence			
9	2822	96.3	2724	6	AK308754	AK308754 Sequence			
10	2571	94.4	2724	9	AF061444	AF061444 Homo sapi			
11	1842	67.6	2082	6	BD135244	BD135244 Novel mam			
C	12	1090	40.0	145165	9	AC078860	AC078860 Homo sapi		
	13	1019	37.4	4095	6	AK075399	AK075399 Homo sapi		
	14	606	22.2	606	6	E42915	E42915 Novel gonad		
	15	606	22.2	606	6	AX016185	AX016185 Sequence		
C	16	212	7.8	162004	9	AC090116	AC090116 Homo sapi		
	17	126	4.6	816	11	BV208561	BV208561 GPR49 209		
	18	60	2.2	60	6	CQ50152	CQ50152 Sequence		
C	19	43	1.6	252477	2	AC136417	AC136417 Rattus no		
	20	43	1.6	252477	2	AC136417	AC136417 Rattus no		

ALIGNMENTS

[illegible]

Db	1441	GCATTTGGAGTGTGTGAAATGCCTATAAGATTTCTTAATCAATGGGAATAAAGGTGACAAC	1500
Qy	1501	AGCAGTATGGACGACCTTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
Db	1501	AGCAGTATGGACGACCTTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
Qy	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGTCAG	1620
Db	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGTCAG	1620
Qy	1621	TGTTTCACTTCCCGACGSCCCTTCAAAACCTGTGAAACCTGCTGATGCTCGCTGATC	1680
Db	1621	TGTTTCACTTCCCGACGSCCCTTCAAAACCTGTGAAACCTGCTGATGCTCGCTGATC	1680
Qy	1681	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTTGTAATGCTTTGGTGACTTCA	1740
Db	1681	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTTGTAATGCTTTGGTGACTTCA	1740
Qy	1741	ACAGTTTTTCAGATCCCTCTGTGACATTTTCCCCCAATTAAACTGTGTAATGTTGGGTCA	1800
Db	1741	ACAGTTTTTCAGATCCCTCTGTGACATTTTCCCCCAATTAAACTGTGTAATGTTGGGTCA	1800
Qy	1801	GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGGCTGGTGGATGGTTCACCT	1860
Db	1801	GCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGGCTGGTGGATGGTTCACCT	1860
Qy	1861	TTTGGCAGCTTTGCAACATGCTGCTGGTGGAGAAATGGGCTGGTTCGCCATGTCAATT	1920
Db	1861	TTTGGCAGCTTTGCAACATGCTGCTGGTGGAGAAATGGGCTGGTTCGCCATGTCAATT	1920
Qy	1921	GGTTTTTGTGCCAATTTTGTCTCAGAAATCAATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980
Db	1921	GGTTTTTGTGCCAATTTTGTCTCAGAAATCAATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980
Qy	1981	GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC	2040
Db	1981	GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC	2040
Qy	2041	CTGAAAGTAATCAATTTTGTCTCTGTGCCCTGTGCCCTTGACCATGGCCGAGTTCCCTCG	2100
Db	2041	CTGAAAGTAATCAATTTTGTCTCTGTGCCCTGTGCCCTTGACCATGGCCGAGTTCCCTCG	2100
Qy	2101	CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGTGCTTGTGCTTTTGGGGAGCCC	2160
Db	2101	CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGTGCTTGTGCTTTTGGGGAGCCC	2160
Qy	2161	AGCACCATGGCTACATGCTCGCTCTCAATCTTGCTCAATTCCTTTGCTTCCTCATGATG	2220
Db	2161	AGCACCATGGCTACATGCTCGCTCTCAATCTTGCTCAATTCCTTTGCTTCCTCATGATG	2220
Qy	2221	ACCAATTGCCCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAAATATTTGG	2280
Db	2221	ACCAATTGCCCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAAATATTTGG	2280
Qy	2281	GACTGCTCTATGTTAAACAACATTTGCCCTGTGCTCTTCAACCAATGCATCTCTAAACTGC	2340
Db	2281	GACTGCTCTATGTTAAACAACATTTGCCCTGTGCTCTTCAACCAATGCATCTCTAAACTGC	2340
Qy	2341	CCTGTGGCTTTCTTGCTCTCTCTTTTAATAAAACCTTTACATTTATCATGCTCTGAAGTA	2400
Db	2341	CCTGTGGCTTTCTTGCTCTCTCTTTTAATAAAACCTTTACATTTATCATGCTCTGAAGTA	2400
Qy	2401	ATTAAAGTTTATCTTCTGTGGTGGTAGTCCCACTTCTCTGCAATGCTCAATCCCCCTCTCTAC	2460
Db	2401	ATTAAAGTTTATCTTCTGTGGTGGTAGTCCCACTTCTCTGCAATGCTCAATCCCCCTCTCTAC	2460
Qy	2461	ATCTTGTCTCAATCCTCTCACTTTTAAGGAGGATCTGGTGAGCCTCGAAGAACCAACCTACGTC	2520
Db	2461	ATCTTGTCTCAATCCTCTCACTTTTAAGGAGGATCTGGTGAGCCTCGAAGAACCAACCTACGTC	2520
Qy	2521	TGGACAAGATCAAAAACACCCAGCTTGATGTCAATTTAACTCTGTATGATGTGAAAACAG	2580
Db	2521	TGGACAAGATCAAAAACACCCAGCTTGATGTCAATTTAACTCTGTATGATGTGAAAACAG	2580

QY	541	GCTTTTAGAAGTTATCGGCATTGCAAGCATGACCTTGGCCCTGAACAAATATACCCAC	600
Db	589	GCTTTTAGAAGTTATCGGCATTGCAAGCATGACCTTGGCCCTGAACAAATATACCCAC	648
QY	601	ATACCAGACTATGCTCTTGGAAACCTCTCCAGCTTGGTAGTCTTACACTCTCCATAACAAT	660
Db	649	ATACCAGACTATGCTCTTGGAAACCTCTCCAGCTTGGTAGTCTTACACTCTCCATAACAAT	708
QY	661	AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
Db	709	AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	768
QY	721	TTAAATTAACAATAACCTTGGTAATTCCTCCACTGCAATTAGACACTCTCCAACTTAAA	780
Db	769	TTAAATTAACAATAACCTTGGTAATTCCTCCACTGCAATTAGACACTCTCCAACTTAAA	828
QY	781	GAACTAGGATTTATAGCAACAAATATCAGGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC	840
Db	829	GAACTAGGATTTATAGCAACAAATATCAGGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC	888
QY	841	CCTTCTCTTATTAACAATATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT	900
Db	889	CCTTCTCTTATTAACAATATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT	948
QY	901	TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGGTGCCTCACAATTAACCTGAA	960
Db	949	TTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGGTGCCTCACAATTAACCTGAA	1008
QY	961	TTTTCCTGATTTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTTAACTGGAGCACAGATC	1020
Db	1009	TTTTCCTGATTTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTTAACTGGAGCACAGATC	1068
QY	1021	TCATCTCTTCCCTCAAACTGTCGAATCAGTTTACCTAAATCTCAAGTCTGATCTGCT	1080
Db	1069	TCATCTCTTCCCTCAAACTGTCGAATCAGTTTACCTAAATCTCAAGTCTGATCTGCT	1128
QY	1081	TACAACCTTATAGAGATTTACCCAGTCTTTCAGTCTGCGAAGCTTCAGAAATTTGAC	1140
Db	1129	TACAACCTTATAGAGATTTACCCAGTCTTTCAGTCTGCGAAGCTTCAGAAATTTGAC	1188
QY	1141	CTAAGACATTAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTCTAGCCTC	1200
Db	1189	CTAAGACATTAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTGTCTAGCCTC	1248
QY	1201	CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTTCAACCCCAATTTTCCACT	1260
Db	1249	CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTTCAACCCCAATTTTCCACT	1308
QY	1261	TTGCCATCCCTAATAAGCTGGACCTATGCTCCAACTCTGCTGCTTTTCCCTATAACT	1320
Db	1309	TTGCCATCCCTAATAAGCTGGACCTATGCTCCAACTCTGCTGCTTTTCCCTATAACT	1368
QY	1321	GGGTTACATGGTTTAACTCACCTTAAATTTAACAGGAAATCATGCTTACAGAGCTTTGATA	1380
Db	1369	GGGTTACATGGTTTAACTCACCTTAAATTTAACAGGAAATCATGCTTACAGAGCTTTGATA	1428
QY	1381	TCATCTGAATACTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACAGTGTGT	1440
Db	1429	TCATCTGAATACTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACAGTGTGT	1488
QY	1441	GCATTTGGAGTGTGAGATGCCCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAAAC	1500
Db	1489	GCATTTGGAGTGTGAGATGCCCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAAAC	1548
QY	1501	AGCAGTATGGAGCACTTCATAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1560
Db	1549	AGCAGTATGGAGCACTTCATAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1608
QY	1561	GACCTTGAAGATTTCTGCTTGATCTTTTGGAGGAGACCTTGAAAGCCCTTCATTTCACTGAG	1620
Db	1609	GACCTTGAAGATTTCTGCTTGATCTTTTGGAGGAGACCTTGAAAGCCCTTCATTTCACTGAG	1668

QY	1621	TGTTACCTTTCCAGGCCCTTTCAAACTCTGTGAACACCTGTCTGTATGGCTGCTGATC	1680
Db	1669	TGTTACCTTTCCAGGCCCTTTCAAACTCTGTGAACACCTGTCTGTATGGCTGCTGATC	1728
QY	1681	AGAAATGGAGTGTGGACCATAGCAGATTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Db	1729	AGAAATGGAGTGTGGACCATAGCAGATTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1788
QY	1741	ACAGTTTTCAGATCCCTCTGTATACATTTTCCCACTTAAACCTGTTAAATTTGGGGTCAATCGCA	1800
Db	1789	ACAGTTTTCAGATCCCTCTGTATACATTTTCCCACTTAAACCTGTTAAATTTGGGGTCAATCGCA	1848
QY	1801	GCAGTGAACATGCTTCACGGGAGTCTCCAGTCCCGTGTGCTGTGTGTGAATGCGTTTCACT	1860
Db	1849	GCAGTGAACATGCTTCACGGGAGTCTCCAGTCCCGTGTGCTGTGTGTGAATGCGTTTCACT	1908
QY	1861	TTTGGCAGCTTTTGCACGACATGCTGTGGGAGAAATGGGGTTGGTTCGCAATGTCATT	1920
Db	1909	TTTGGCAGCTTTTGCACGACATGCTGTGGGAGAAATGGGGTTGGTTCGCAATGTCATT	1968
QY	1921	GGTTTTCCTCCATTTTTCCTTTCAGAAATCATCTGTTTTCCTGCTTACTCTGCGCAGCCCTG	1980
Db	1969	GGTTTTCCTCCATTTTTCCTTTCAGAAATCATCTGTTTTCCTGCTTACTCTGCGCAGCCCTG	2028
QY	1981	GAGCGTGGTTCCTGTGAAATATTTCTGCAAAATTTTGAACAGAAAGCTCCATTTTCTAGC	2040
Db	2029	GAGCGTGGTTCCTGTGAAATATTTCTGCAAAATTTTGAACAGAAAGCTCCATTTTCTAGC	2088
QY	2041	CTGAAAGTAATCAATTTTTCCTTCTGTGCTGTGGCCTTGGACCATGCGCGCAGTTTCCCTG	2100
Db	2089	CTGAAAGTAATCAATTTTTCCTTCTGTGCTGTGGCCTTGGACCATGCGCGCAGTTTCCCTG	2148
QY	2101	CTGGGTGGCAGCAGTATGGGCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGGCC	2160
Db	2149	CTGGGTGGCAGCAGTATGGGCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGGAGGCC	2208
QY	2161	AGCACCATGGCTACATGCTGCTCTCATCTTGTCAATTTCCCTTTCCTTCTCATGATG	2220
Db	2209	AGCACCATGGCTACATGCTGCTCTCATCTTGTCAATTTCCCTTTCCTTCTCATGATG	2268
QY	2221	ACCATGGCTACACCAAGCTCTACTGCAATTTTGAACAGGGAGACCTTGGAGAAATTTTGG	2280
Db	2269	ACCATGGCTACACCAAGCTCTACTGCAATTTTGAACAGGGAGACCTTGGAGAAATTTTGG	2328
QY	2281	GACTGCTCTATGTTAAACACATTTGCTGCTCTTTCACCACTGCTCCTTAAACTGC	2340
Db	2329	GACTGCTCTATGTTAAACACATTTGCTGCTCTTTCACCACTGCTCCTTAAACTGC	2388
QY	2341	CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTGAAAGTA	2400
Db	2389	CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTGAAAGTA	2448
QY	2401	ATTAAGTTTATCTTCTTGTGGTGTAGTCCCACTTCTGCTGCTCAATTCCTTCTCTAC	2460
Db	2449	ATTAAGTTTATCTTCTTGTGGTGTAGTCCCACTTCTGCTGCTCAATTCCTTCTCTAC	2508
QY	2461	ATCTGTTCATCTCTCACTTTTAAAGGAGATCTGGTGAGCCTTGAGAAAGCAACCTACGTC	2520
Db	2509	ATCTGTTCATCTCTCACTTTTAAAGGAGATCTGGTGAGCCTTGAGAAAGCAACCTACGTC	2568
QY	2521	TGGACAAGATCAAAACACCCCAAGCTTGTATGTCAATTAACCTCTGATGATGTGCAAAAAACAG	2580
Db	2569	TGGACAAGATCAAAACACCCCAAGCTTGTATGTCAATTAACCTCTGATGATGTGCAAAAAACAG	2628
QY	2581	TCCTGTGATCAACTCAAGCCTTGTGTAACCTTTTACAGCTCCAGCATCATTATGACCTG	2640
Db	2629	TCCTGTGATCAACTCAAGCCTTGTGTAACCTTTTACAGCTCCAGCATCATTATGACCTG	2688
QY	2641	CCTCCAGTTCCTGGCCATCACAGCTTATCCAGTGTAGTGTGCACTTCTTCTCTCT	2700
Db	2689	CCTCCAGTTCCTGGCCATCACAGCTTATCCAGTGTAGTGTGCACTTCTTCTCTCT	2748
QY	2701	GTGGCAATTTGTCCCATGTCTCTAA	2724

Qy	13261	TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCCTGTCGTCTTTCCCTAAACT	13262	TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCCTGTCGTCTTTCCCTAAACT	13263	TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCCTGTCGTCTTTCCCTAAACT	13264	TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCCTGTCGTCTTTCCCTAAACT
Db	13309	TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCCTGTCGTCTTTCCCTAAACT	13310	TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCCTGTCGTCTTTCCCTAAACT	13311	TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCCTGTCGTCTTTCCCTAAACT	13312	TTGGCATCCCTAATAAAGCTGGACCTATCGTCCAAACCTCCTGTCGTCTTTCCCTAAACT
Qy	13321	GGGTTACATGGTTTAACTCACTTAAATAATACAGGAAATCATGCCCTTACAGAGCTTGATA	13322	GGGTTACATGGTTTAACTCACTTAAATAATACAGGAAATCATGCCCTTACAGAGCTTGATA	13323	GGGTTACATGGTTTAACTCACTTAAATAATACAGGAAATCATGCCCTTACAGAGCTTGATA	13324	GGGTTACATGGTTTAACTCACTTAAATAATACAGGAAATCATGCCCTTACAGAGCTTGATA
Db	13369	GGGTTACATGGTTTAACTCACTTAAATAATACAGGAAATCATGCCCTTACAGAGCTTGATA	13370	GGGTTACATGGTTTAACTCACTTAAATAATACAGGAAATCATGCCCTTACAGAGCTTGATA	13371	GGGTTACATGGTTTAACTCACTTAAATAATACAGGAAATCATGCCCTTACAGAGCTTGATA	13372	GGGTTACATGGTTTAACTCACTTAAATAATACAGGAAATCATGCCCTTACAGAGCTTGATA
Qy	13381	TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACACAGTCTGT	13382	TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACACAGTCTGT	13383	TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACACAGTCTGT	13384	TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACACAGTCTGT
Db	1429	TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACACAGTCTGT	1430	TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACACAGTCTGT	1431	TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACACAGTCTGT	1432	TCATCTGAAACCTTTCCAGAACTCAAGGTTTATAGAAATGCTTATGCTTACACAGTCTGT
Qy	1441	GCATTTGGAGTGTGTGAGAAATGCCCTAATAAGATTTCTAATCAATGGAAATAAAGGTGACAAC	1442	GCATTTGGAGTGTGTGAGAAATGCCCTAATAAGATTTCTAATCAATGGAAATAAAGGTGACAAC	1443	GCATTTGGAGTGTGTGAGAAATGCCCTAATAAGATTTCTAATCAATGGAAATAAAGGTGACAAC	1444	GCATTTGGAGTGTGTGAGAAATGCCCTAATAAGATTTCTAATCAATGGAAATAAAGGTGACAAC
Db	1489	GCATTTGGAGTGTGTGAGAAATGCCCTAATAAGATTTCTAATCAATGGAAATAAAGGTGACAAC	1490	GCATTTGGAGTGTGTGAGAAATGCCCTAATAAGATTTCTAATCAATGGAAATAAAGGTGACAAC	1491	GCATTTGGAGTGTGTGAGAAATGCCCTAATAAGATTTCTAATCAATGGAAATAAAGGTGACAAC	1492	GCATTTGGAGTGTGTGAGAAATGCCCTAATAAGATTTCTAATCAATGGAAATAAAGGTGACAAC
Qy	1501	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1502	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1503	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1504	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT
Db	1549	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1550	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1551	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1552	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT
Qy	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGAAAGCCCTTCATTTCAGTGCAG	1562	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGAAAGCCCTTCATTTCAGTGCAG	1563	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGAAAGCCCTTCATTTCAGTGCAG	1564	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGAAAGCCCTTCATTTCAGTGCAG
Db	1609	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGAAAGCCCTTCATTTCAGTGCAG	1610	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGAAAGCCCTTCATTTCAGTGCAG	1611	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGAAAGCCCTTCATTTCAGTGCAG	1612	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAGACCTGAAAGCCCTTCATTTCAGTGCAG
Qy	1621	TGTTTCACCTCCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTCGCTGATC	1622	TGTTTCACCTCCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTCGCTGATC	1623	TGTTTCACCTCCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTCGCTGATC	1624	TGTTTCACCTCCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTCGCTGATC
Db	1669	TGTTTCACCTCCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTCGCTGATC	1670	TGTTTCACCTCCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTCGCTGATC	1671	TGTTTCACCTCCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTCGCTGATC	1672	TGTTTCACCTCCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTCGCTGATC
Qy	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA	1682	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA	1683	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA	1684	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA
Db	1729	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA	1730	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA	1731	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA	1732	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACCTTACTTGTAAATGCTTTGGTGACTTCA
Qy	1741	ACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAAATGGGGTCATCGCA	1742	ACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAAATGGGGTCATCGCA	1743	ACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAAATGGGGTCATCGCA	1744	ACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAAATGGGGTCATCGCA
Db	1789	ACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAAATGGGGTCATCGCA	1790	ACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAAATGGGGTCATCGCA	1791	ACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAAATGGGGTCATCGCA	1792	ACAGTTTTTCAGATCCCTCTGTACATTTCCGCCATTTAAACTGTGTAAATGGGGTCATCGCA
Qy	1801	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGCTGCTGTGTGTGAATGCGTTCACT	1802	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGCTGCTGTGTGTGAATGCGTTCACT	1803	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGCTGCTGTGTGTGAATGCGTTCACT	1804	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGCTGCTGTGTGTGAATGCGTTCACT
Db	1849	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGCTGCTGTGTGTGAATGCGTTCACT	1850	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGCTGCTGTGTGTGAATGCGTTCACT	1851	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGCTGCTGTGTGTGAATGCGTTCACT	1852	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGCTGCTGTGTGTGAATGCGTTCACT
Qy	1861	TTTGGCAGCTTTTGACGACATGCTGCTGGTGGAGAAATGGGTGTGGTTCGCCATGCAATT	1862	TTTGGCAGCTTTTGACGACATGCTGCTGGTGGAGAAATGGGTGTGGTTCGCCATGCAATT	1863	TTTGGCAGCTTTTGACGACATGCTGCTGGTGGAGAAATGGGTGTGGTTCGCCATGCAATT	1864	TTTGGCAGCTTTTGACGACATGCTG

[illegible]

Qy	1361	TTGGCATCCCTTAATAAAGCTGGGACCTATCGTCCAACCTCCTGTGCTCTTTTTCCTATAAACT	1321
Db	1522	TTGGCATCCCTTAATAAAGCTGGGACCTATCGTCCAACCTCCTGTGCTCTTTTTCCTATAAACT	1581
Qy	1321	GGGTTACATGTTTTAACTCACTTAAAATTAAACAGGAATCATGCCTTTACAGAGCTTGATA	1380
Db	1582	GGGTTACATGTTTTAACTCACTTAAAATTAAACAGGAANTANGGCCITACAGAGCTTGATA	1641
Qy	1381	TCACTCTGAAAACTTTTCCAGAACCTCAAGGTTATAGAAATGCTTATGTTACCAGTGCCTGT	1440
Db	1642	TCATCTGAAAACTTTTCCAGAACCTCAAGGTTATAGAAATGCTTATGTTACCAGTGCCTGT	1701
Qy	1441	GCAATTTGGAGTGTGAGAATGCCCCTATAAGATTTCTAATCAATGGAATAAAGTGACAAC	1500
Db	1702	GCAATTTGGAGTGTGAGAATGCCCCTATAAGATTTCTAATCAATGGAATAAAGTGACAAC	1761
Qy	1501	AGCAGTATGACGACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1560
Db	1762	AGCAGTATGACGACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1821
Qy	1561	GACCTTGAAGATTTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGCAG	1620
Db	1822	GACCTTGAAGATTTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGCAG	1881
Qy	1621	TGTTTCACTTCCCAGGCCCCCTTCAAACCCCTGTGAAACACTGCTGTGAATGCTGGCTGATC	1680
Db	1882	TGTTTCACTTCCCAGGCCCCCTTCAAACCCCTGTGAAACACTGCTGTGAATGCTGGCTGATC	1941
Qy	1681	AGAAATGGAGTGGACCATAGCAGTTCTGGCACCTTACTTTGTAATGCTTTTGGTGACTTCA	1740
Db	1942	AGAAATGGAGTGGACCATAGCAGTTCTGGCACCTTACTTTGTAATGCTTTTGGTGACTTCA	2001
Qy	1741	ACAGTTTTTCAGATCCCCCTCTGTACATTTTCCCCCAATAAACTGTTAATTTGGGGTCATCGCA	1800
Db	2002	ACAGTTTTTCAGATCCCCCTCTGTACATTTTCCCCCAATAAACTGTTAATTTGGGGTCATCGCA	2061
Qy	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGTGCTGGTGTGGAATGCGTTCACT	1860
Db	2062	GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGTGCTGGTGTGGAATGCGTTCACT	2121
Qy	1861	TTTTGGCAGCTTTGCACACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAATT	1920
Db	2122	TTTTGGCAGCTTTGCACACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAATT	2181
Qy	1921	GGTTTTTTTGTCATTTTTTGCTTCAGAAATCATCTGTTTTCTTGCTTACTCTGGCAGCCCTG	1980
Db	2182	GGTTTTTTTGTCATTTTTTGCTTCAGAAATCATCTGTTTTCTTGCTTACTCTGGCAGCCCTG	2241
Qy	1981	GAGCGTGGGTTCTCTGTAATAATFTCTGAAAAATTTGAAAAAGCTCCATTTTCTPAGC	2040
Db	2242	GAGCGTGGGTTCTCTGTAATAATFTCTGAAAAATTTGAAAAAGCTCCATTTTCTPAGC	2301
Qy	2041	CTGAAAGTAAATCATTTTGTCTCTGTCCTCTGTCCTGCTTGGACCATGGCCGAGTTCGCCCTG	2100
Db	2302	CTGAAAGTAAATCATTTTGTCTCTGTCCTCTGTCCTGCTTGGACCATGGCCGAGTTCGCCCTG	2361
Qy	2101	CTGGTGGCAGCAGTAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGGCTTTTGGGAGCCCC	2160
Db	2362	CTGGTGGCAGCAGTAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGGCTTTTGGGAGCCCC	2421
Qy	2161	AGCACCATGGGCTACATGGTGCCTCTCATCTGTCTCAATTTCCCTTTGCTTCTCATGATG	2220
Db	2422	AGCACCATGGGCTACATGGTGCCTCTCATCTGTCTCAATTTCCCTTTGCTTCTCATGATG	2481
Qy	2221	ACCAATGCTTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTTGG	2280
Db	2482	ACCAATGCTTACACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTTGG	2541
Qy	2281	GACTGCTCTATGGTAAACACATGGCCCTGTGCTCTTTCACCAACTGCATCTCTAAACTGC	2340
Db	2542	GACTGCTCTATGGTAAACACATGGCCCTGTGCTCTTTCACCAACTGCATCTCTAAACTGC	2601
Qy	2341	CCTGTGCTCTTCTGTGCTCTCTCTCTTTTAAATAAACCTTACATTTTATFACGTCTCTGAAGTA	2400

[illegible][illegible]

QY 301 GCTCTGACATATTTCCCAAGGAGCATCTCACTGGSCCTTTACAGTCTTTAAAGTTCTTATG 360
DB 301 GCTCTGACATATTTCCCAAGGAGCATCTCACTGGSCCTTTACAGTCTTTAAAGTTCTTATG 360
QY 361 CTGCAGAAATAATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCCTT 420
DB 361 CTGCAGAAATAATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCCTT 420
QY 421 CAATCCCTGCGTCTGGATGCTTAAACAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
DB 421 CAATCCCTGCGTCTGGATGCTTAAACAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
QY 481 CTGCATTCCTCGAGGACCTGCTGGCTGATGACAAATGCTTAAACAGAAATCCCCTGCCAG 540
DB 481 CTGCATTCCTCGAGGACCTGCTGGCTGATGACAAATGCTTAAACAGAAATCCCCTGCCAG 540
QY 541 GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
DB 541 GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
QY 601 ATACAGACTATGCTTTGGAAAACCTCTCCAGCTTGGTAGTCTACATCTCGATAACAAT 660
DB 601 ATACAGACTATGCTTTGGAAAACCTCTCCAGCTTGGTAGTCTACATCTCGATAACAAT 660
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
DB 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT 720
QY 721 TTAATAATACAATAAATCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAAACCTTAAA 780
DB 721 TTAATAATACAATAAATCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAAACCTTAAA 780
QY 781 GAACTAGAGATTTATAGCAACAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB 781 GAACTAGAGATTTATAGCAACAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
QY 841 CCTCTCTTATTAACAATAATTTCTATGCAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
DB 841 CCTCTCTTATTAACAATAATTTCTATGCAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
QY 901 TTTCAACATTTACTGGAATTAAGAACACTGACTCTGTAATGGTGGCTCAAAAATTAAGTAA 960
DB 901 TTTCAACATTTACTGGAATTAAGAACACTGACTCTGTAATGGTGGCTCAAAAATTAAGTAA 960
QY 961 TTTCTGATTTAACTGGAAGTCAAAACCTGGAGTCTGACTTTAACTGGAGCACAGATC 1020
DB 961 TTTCTGATTTAACTGGAAGTCAAAACCTGGAGTCTGACTTTAACTGGAGCACAGATC 1020
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAAGTCTAGATCTGCT 1080
DB 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAAGTCTAGATCTGCT 1080
QY 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTCAC 1140
DB 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTCAC 1140
QY 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTCGTTAGCCTC 1200
DB 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTCGTTAGCCTC 1200
QY 1201 CGATCGCTGAATTTGGCTTTGGAACAAAATTTGCTATTATTTACCCCAATGCAATTTTCCACT 1260
DB 1201 CGATCGCTGAATTTGGCTTTGGAACAAAATTTGCTATTATTTACCCCAATGCAATTTTCCACT 1260
QY 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAAACCTCTGCTGCTTTTCTCTATAACT 1320
DB 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAAACCTCTGCTGCTTTTCTCTATAACT 1320
QY 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTTGATA 1380
DB 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTTGATA 1380
QY 1381 TCATCTGAAACCTTTCCGAACTCAAGGTTATAGAAATGSCCTTATGCTTACCAGTCTGT 1440

DB 1381 TCATCTGAAACCTTTCCGAACTCAAGGTTATAGAAATGSCCTTATGCTTACCAGTCTGT 1440
QY 1441 GCATTTGGAGTGTGTGAGATGCCCTATAAGATTTCTTAATCAATGGAATAAAGGTGACAAAC 1500
DB 1441 GCATTTGGAGTGTGTGAGATGCCCTATAAGATTTCTTAATCAATGGAATAAAGGTGACAAAC 1500
QY 1501 AGCAGTATGAGACGACTTCAATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
DB 1501 AGCAGTATGAGACGACTTCAATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
QY 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAACCTGAAAGCCCTTCAATTCAGTGGCAG 1620
DB 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAACCTGAAAGCCCTTCAATTCAGTGGCAG 1620
QY 1621 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTTGAACACCTGTTGATGGCTGGCTGATC 1680
DB 1621 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTTGAACACCTGTTGATGGCTGGCTGATC 1680
QY 1681 AGAATTCGAGTGTGGACCATAGCAGTTCTGSCACTTACTTGTAAATGCTTTGCTGACTTCA 1740
DB 1681 AGAATTCGAGTGTGGACCATAGCAGTTCTGSCACTTACTTGTAAATGCTTTGCTGACTTCA 1740
QY 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTCAATGCA 1800
DB 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTCAATGCA 1800
QY 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTGGCTGCTGGCTGGTGGATGCGTTCAT 1860
DB 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTGGCTGCTGGCTGGTGGATGCGTTCAT 1860
QY 1861 TTTGCGAGCTTTGACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGCAATGTCATT 1920
DB 1861 TTTGCGAGCTTTGACGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGCAATGTCATT 1920
QY 1921 GGTTTTTGTGCTTTCAGAAATCACTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
DB 1921 GGTTTTTGTGCTTTCAGAAATCACTGTTTTCTGCTTACTCTGGCAGCCCTG 1980
QY 1981 GAGCGTGGTTCCTGTGAAATATCTGCAAAAATTTGAAAGCAAGCTCCATTTCTTAGC 2040
DB 1981 GAGCGTGGTTCCTGTGAAATATCTGCAAAAATTTGAAAGCAAGCTCCATTTCTTAGC 2040
QY 2041 CTGAAAGTAAATCAATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 2041 CTGAAAGTAAATCAATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
DB 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
QY 2161 AGCACCATGGCTACATGGTGGCTCTCATCTGCTCAATTCCTTGGCTTCCCTCATGATG 2220
DB 2161 AGCACCATGGCTACATGGTGGCTCTCATCTGCTCAATTCCTTGGCTTCCCTCATGATG 2220
QY 2221 ACCATTCCTACACCAAGCTCTACTGCAATTTGACAAAGGAGACCTGGAGAAATTTGG 2280
DB 2221 ACCATTCCTACACCAAGCTCTACTGCAATTTGACAAAGGAGACCTGGAGAAATTTGG 2280
QY 2281 GACTGCTCTATGTTAAACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
DB 2281 GACTGCTCTATGTTAAACACATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
QY 2341 CTGCTGGCTTTCTGCTGCTTCTCTTATTAACCTTACATTTATCAGTCCCTGGAAGTA 2400
DB 2341 CTGCTGGCTTTCTGCTGCTTCTCTTATTAACCTTACATTTATCAGTCCCTGGAAGTA 2400
QY 2401 ATTAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
DB 2401 ATTAAGTTTATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
QY 2461 ATCTTGTTCATCTCTCTTAAAGGAGATCTGTTAGCTGCTGAGAAAGCAAACTTACCTGCTC 2520

QY 961 TTCTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1020
Db 961 TTCTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTTAACTGGAGCAGATC 1020
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
Db 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
QY 1081 TACAACCTATTAGAAAGATTACCCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db 1081 TACAACCTATTAGAAAGATTACCCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
QY 1141 CTAAGACATATGAATCTACGAATTTAAAGTTGACACTTTTCCAGCAGTTCCTAGCCTC 1200
Db 1141 CTAAGACATATGAATCTACGAATTTAAAGTTGACACTTTTCCAGCAGTTCCTAGCCTC 1200
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTACCCCAATGCAATTTTCCACT 1260
QY 1261 TTGCCATCCCTAATAAAGCTGCACTATCGTCCAAACCTCTGTCTGCTTTTCCCTATAACT 1320
Db 1261 TTGCCATCCCTAATAAAGCTGCACTATCGTCCAAACCTCTGTCTGCTTTTCCCTATAACT 1320
QY 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACAGTGTGT 1440
Db 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACAGTGTGT 1440
QY 1441 GCATTTGGAGTGTGAGAAATGCTTATAGAAATTTCTAATCAATGAAATPAAAGGTGACAA 1500
Db 1441 GCATTTGGAGTGTGAGAAATGCTTATAGAAATTTCTAATCAATGAAATPAAAGGTGACAA 1500
QY 1501 AGCAGTATGGAGGACCTTCATAGAAAGTGTGGAATGTTTCAGGCTCAAGATGAAGT 1560
Db 1501 AGCAGTATGGAGGACCTTCATAGAAAGTGTGGAATGTTTCAGGCTCAAGATGAAGT 1560
QY 1561 GACCTTGAAGATTTCCTGCTTACCTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db 1561 GACCTTGAAGATTTCCTGCTTACCTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
QY 1621 TGTTCACCTTCCAGGCCCCCTTCAAAACCTGTGAAACACCTGCTGATGCGTGCATC 1680
Db 1621 TGTTCACCTTCCAGGCCCCCTTCAAAACCTGTGAAACACCTGCTGATGCGTGCATC 1680
QY 1681 AGAATTTGGAGTGTGACCATAGCAGTCTTGGCACTTACTTGTAAATGCTTTGGTGAATCA 1740
Db 1681 AGAATTTGGAGTGTGACCATAGCAGTCTTGGCACTTACTTGTAAATGCTTTGGTGAATCA 1740
QY 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCAATTAACCTGTTAAATTTGGGCTCATCGCA 1800
Db 1741 ACAGTTTTCAGATCCCTCTGACATTTCCCCCAATTAACCTGTTAAATTTGGGCTCATCGCA 1800
QY 1801 CGAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGCTGGTGTGATGCGTTCAC 1860
Db 1801 CGAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGCTGGTGTGATGCGTTCAC 1860
QY 1861 TTTGGCAGCTTGCACGACATGGTCCCTGCTGGGAGAAATGGGTTGGTGGCCTGCAAT 1920
Db 1861 TTTGGCAGCTTGCACGACATGGTCCCTGCTGGGAGAAATGGGTTGGTGGCCTGCAAT 1920
QY 1921 GGTGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTTCCCTGCTTACTCTGGCAGCCCTG 1980
Db 1921 GGTGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTTCCCTGCTTACTCTGGCAGCCCTG 1980
QY 1981 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
QY 2041 CTGAAAGTAATCAATTTTGTCTGTGCCCTGTGCTGGCCTTGACCATGGCGCAGTTCCCTCTG 2100

Db 2041 CTGAAAGTAATCAATTTTGTCTGTGCCCTGTGCTGCTTGCATGGCGCAGTTCCCTCTG 2100
QY 2101 CTGGGTGGCAGCAAGTATGCGCTCCCTCTCTGCTGCTGCTTTGGGAGAGCC 2160
Db 2101 CTGGGTGGCAGCAAGTATGCGCTCCCTCTCTGCTGCTGCTTTGGGAGAGCC 2160
QY 2161 AGCACCATGGGTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTGGTTCCTCATGATG 2220
Db 2161 AGCACCATGGGTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTGGTTCCTCATGATG 2220
QY 2221 ACCATTTGCTACACCAAGCTCTACTGCAATTTGACAAGGAGAGACCTGGAGAAATATTGG 2280
Db 2221 ACCATTTGCTACACCAAGCTCTACTGCAATTTGACAAGGAGAGACCTGGAGAAATATTGG 2280
QY 2281 GACTGCTCTATGTTAAACACATTTGCCCTGTGTCTTTCACCAACCTGATCCTAAACCTGC 2340
Db 2281 GACTGCTCTATGTTAAACACATTTGCCCTGTGTCTTTCACCAACCTGATCCTAAACCTGC 2340
QY 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCCCTGAAGTA 2400
Db 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCCCTGAAGTA 2400
QY 2401 ATTAAGTTTATCTCTTGTGTTAGTCCCACTCTCTGCTGCTCTCAATCCCTTCTCTAC 2460
Db 2401 ATTAAGTTTATCTCTTGTGTTAGTCCCACTCTCTGCTGCTCTCAATCCCTTCTCTAC 2460
QY 2461 ATCTTTGTTCAATCTCACTTTAAGGAGATCTGTGAGCCTGAGAAAGCAACCTACGTC 2520
Db 2461 ATCTTTGTTCAATCTCACTTTAAGGAGATCTGTGAGCCTGAGAAAGCAACCTACGTC 2520
QY 2521 TGGCAAGATCAAAACACCCCAAGCTTGAATGATGATCAATTAACCTGATGATGTCGAAAAACAG 2580
Db 2521 TGGCAAGATCAAAACACCCCAAGCTTGAATGATGATCAATTAACCTGATGATGTCGAAAAACAG 2580
QY 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACCAGCTCCAGCATCACCTTATGACCTG 2640
Db 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACCAGCTCCAGCATCACCTTATGACCTG 2640
QY 2641 CCTCCCAGTTCGGTGCATACACAGCTTATCCAGTGAATCCAGTGAAGTGCATCTTCTCT 2700
Db 2641 CCTCCCAGTTCGGTGCATACACAGCTTATCCAGTGAATCCAGTGAAGTGCATCTTCTCTCT 2700
QY 2701 GTGGCAATTTGCCCATGCTCTAA 2724
Db 2701 GTGGCAATTTGCCCATGCTCTAA 2724

RESULT 11
BD135244
LOCUS
DEFINITION Novel mammalian G protein-coupled receptor having extracellular
leucine-rich repeating domain.
ACCESSION
BD135244
VERSION
BD135244.1 GI:23230189
KEYWORDS
JP 2002507406-A/2.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1 (bases 1 to 2082)
AUTHORS
Hsueh, A. J. W.; Hsu, S. Y.; Liang, S. G. and Spe, P. J. V. D.
TITLE
Novel mammalian G protein-coupled receptor having extracellular
leucine-rich repeating domain
JOURNAL
Patent: JP 2002507406-A 2 12-MAR-2002;
THE BOARD OF TRUSTEES OF THE ILELAND STANFORD JUNIOR UNIVERSITY,
AKZO NOBEL NV
COMMENT
OS Homo sapiens (human)
PN JP 2002507406-A/2
PD 12-MAR-2002
PF 25-MAR-1999 JP 2000537903
PR 26-MAR-1998 US 60/079501
PI AARON J W HSUEH, SHEAU YU HSU, SHAN GUANG LIANG, PETRUS JOHANNES

PI	VAN DER SPEK	
PC	C12N15/09,A01K67/027,C07K14/705,C07K16/28,C12N1/15,C12N1/19,	
PC	C12N1/21,	
PC	C12N5/10,C12P21/08,C12N15/00,C12N5/00	
CC	Novel mammalian G protein-coupled receptor having CC	
CC	extracellular	
CC	leucine-rich repeating domain	
PH	Key	Location/Qualifiers
FT	source	1..2082
FT	source	/organism="Homo sapiens (human)"
FEATURES		
source		
ORIGIN		
	Query Match	67.6%; Score 1842; DB 6; Length 2082;
	Best Local Similarity	99.9%; Pred. No. 0;
	Matches 1942; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	781	GAAC TAGGATTT CATAGCAACAAATATCAGGTGATACCTGAGAAAGCAATTTGATGCAAC 840
DB	139	GAAC TAGGATTT CATAGCAACAAATATCAGGTGATACCTGAGAAAGCAATTTGATGCAAC 198
QY	841	CTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCATTTGTTGGGAGATCTGCT 900
DB	199	CTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCATTTGTTGGGAGATCTGCT 258
QY	901	TTTCAACATTTTAACTGGAACCTGAAACCTGAGAGTCTGACATTTAACTGGAGACACAGATC 960
DB	259	TTTCAACATTTTAACTGGAACCTGAAACCTGAGTCTGAAATGGTGGCTCACAATAAATCTGAA 318
QY	961	TTTCTCTGATTTAACTGGAACCTGAAACCTGAGAGTCTGACATTTAACTGGAGACACAGATC 1020
DB	319	TTTCTCTGATTTAACTGGAACCTGAAACCTGAGAGTCTGACATTTAACTGGAGACACAGATC 378
QY	1021	TCATCTCTTCTCAAAACCTCTGCAATCAGTTTACCTTAATCTCCAAAGTGTAGATCTGCT 1080
DB	379	TCATCTCTTCTCAAAACCTCTGCAATCAGTTTACCTTAATCTCCAAAGTGTAGATCTGCT 438
QY	1081	TACACCTATTAGAGATTTACCGAGTTTTCAGTCTGCGAAGCTTCCAGAAATTCAG 1140
DB	439	TACAACTATTAGAGATTTACCGAGTTTTCAGTCTGCGAAGCTTCCAGAAATTCAG 498
QY	1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTCCAGAGTTGCTTAGCCTC 1200
DB	499	CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTCCAGAGTTGCTTAGCCTC 558
QY	1201	CGATCGCTGAATTTGGCTTTGGAAACAAAATTTGCTATTATTTACCCCAATGCAATTTTCCACT 1260
DB	559	CGATCGCTGAATTTGGCTTTGGAAACAAAATTTGCTATTATTTACCCCAATGCAATTTTCCACT 618
QY	1261	TTGGCATCCCTAATAAGCTGACCTATGCTGCAACCTCTGCTGCTTTTCCCTATAACT 1320
DB	619	TTGGCATCCCTAATAAGCTGACCTATGCTGCAACCTCTGCTGCTTTTCCCTATAACT 678
QY	1321	GGGTTACATGGTTTAACTCACATTTAAATTTAAACAGAAATCATGCTTACAGAGCTTGATA 1380
DB	679	GGGTTACATGGTTTAACTCACATTTAAATTTAAACAGAAATCATGCTTACAGAGCTTGATA 738
QY	1381	TCATCTGAAACCTTTCCGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACCAAGTGTGT 1440
DB	739	TCATCTGAAACCTTTCCGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACCAAGTGTGT 798
QY	1441	GCATTTGGAGTGTGAGAAATGCCCTTATAGATTTCTAATCAATGAATGAATGAAGTGAAC 1500
DB	799	GCATTTGGAGTGTGAGAAATGCCCTTATAGATTTCTAATCAATGAATGAATGAAGTGAAC 858
QY	1501	ASCAGTATGGAGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAACGT 1560
DB	859	AGCAGTATGGAGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 918

QY	1561	GACCTTGAAGATTTCTGCTTGACATTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
DB	919	GACCTTGAAGATTTCTGCTTGACATTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 978
QY	1621	TGTTTCAACCTTCCCGAGGCCCTTCAAAACCTGTTGAAACACCTGCTTGTGATGGCTGGCTGATC 1680
DB	979	TGTTTCAACCTTCCCGAGGCCCTTCAAAACCTGTTGAAACACCTGCTTGTGATGGCTGGCTGATC 1038
QY	1681	AGAAATGGAGTGTGACCATAGCAGTTCTTGGCACCTTACTTGTAAATGCTTTTGTGATCTTCA 1740
DB	1039	AGAAATGGAGTGTGACCATAGCAGTTCTTGGCACCTTACTTGTAAATGCTTTTGTGATCTTCA 1098
QY	1741	ACAGTTTTCAGATCCCTCTGTGACATTTTCCCCCATATAAATCTGTTAAATTTGGGGTCAATCGCA 1800
DB	1099	ACAGTTTTCAGATCCCTCTGTGACATTTTCCCCCATATAAATCTGTTAAATTTGGGGTCAATCGCA 1158
QY	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1860
DB	1159	GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1218
QY	1861	TTTGGCAGCTTTTGCACGACATGCTGGTGGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCATT 1920
DB	1219	TTTGGCAGCTTTTGCACGACATGCTGGTGGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCATT 1278
QY	1921	GGTTTTTGTCCATTTTGTCTCAGAAATCACTGTGTTTTCTGCTTACTCTTGGCAGCCCTG 1980
DB	1279	GGTTTTTGTCCATTTTGTCTCAGAAATCACTGTGTTTTCTGCTTACTCTTGGCAGCCCTG 1338
QY	1981	GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB	1339	GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 1398
QY	2041	CTGAAAGTAAATCAATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
DB	1399	CTGAAAGTAAATCAATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1458
QY	2101	CTGGTGGCAGCAAGATATGGGCTCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
DB	1459	CTGGTGGCAGCAAGATATGGGCTCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1518
QY	2161	AGCACCATGGCTACATGCTGCTCATCTGCTCAATTCCTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 2220
DB	1519	AGCACCATGGCTACATGCTGCTCATCTGCTCAATTCCTTCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCT 1578
QY	2221	ACCAATGGCTACACCAAGCTCTACTGCAATTTTGACAAAGGAGAGACCTGGAGAAATATTTGG 2280
DB	1579	ACCAATGGCTACACCAAGCTCTACTGCAATTTTGACAAAGGAGAGACCTGGAGAAATATTTGG 1638
QY	2281	GACTGCTCTATGGTAAACACATTTGCCCTGTTGCTCTTTCACCAACTGCATCTTAAACTGC 2340
DB	1639	GACTGCTCTATGGTAAACACATTTGCCCTGTTGCTCTTTCACCAACTGCATCTTAAACTGC 1698
QY	2341	CCTGTTGGCTTTTGTCT 2400
DB	1699	CCTGTTGGCTTTTGTCT 1758
QY	2401	ATTAAGTTTATCTCTCTGTTGTTAGTCCCACTCTCTGCTGCTCTCAATCCCTCTCTCTCTCTCTCTCTCT 2460
DB	1759	ATTAAGTTTATCTCTCTGTTGTTAGTCCCACTCTCTGCTGCTCTCAATCCCTCTCTCTCTCTCTCTCTCT 1818
QY	2461	ATCTGTTTCAATCT 2520
DB	1819	ATCTGTTTCAATCT 1878
QY	2521	TGGCAAGATCAAAACACCCCAAGCTTGTGATGTTAACTCTCTGATGATGTCGAAAAACAG 2580
DB	1879	TGGCAAGATCAAAACACCCCAAGCTTGTGATGTTAACTCTCTGATGATGTCGAAAAACAG 1938
QY	2581	TCCTGTGACTCAACTCAAGCTTGTGATGTTAACTCTCTGATGATGTCGAAAAACAG 2640
DB	1939	TCCTGTGACTCAACTCAAGCTTGTGATGTTAACTCTCTGATGATGTCGAAAAACAG 1998
QY	2641	CCTCCAGTTCCTGTCATCAACAGCTTTATCCAGTGTGAGAGCTGCCATCTTTCTCTCTCTCTCTCTCTCTCT 2700

|||||
1999 CCTCCAGTTCGTCCTACACAGCTTATCAGTGAAGTCTGAGCTGCCATCTTCTCTCT 2058
|||||
Qy 2701 GTGGCATTGTCCCATGTCTCTAA 2724
|||||
Db 2059 GTGGCATTGTCCCATGTCTCTAA 2082
|||||

AC078860 145165 bp DNA linear PRI 23-JAN-2003
Homo sapiens 12 BAC RP11-186F10 (Roswell Park Cancer Institute
Human BAC Library) complete sequence.
AC078860
AC078860.19 GI:13491193
HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 145165)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Banks,T., Barbaria,J.,
Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,
Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,
Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
Denn,A.L., Ding,X., Dinh,H.H., Douthwaite,K.J., Draper,H.,
Dugan-Rocha,S., Durbin,K.J., Earhart,C., Edgar,D., Edwards,C.C.,
Elhaj,C., Embling,S., Escotto,M., Falls,T., Ferraguto,D.,
Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J.,
Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,
Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,
Hart,M., Haviak,P., Hawes,A., Hernandez,J., Hernandez,O.,
Hodgson,A., Hogue,M., Holloway,C., Hollins,B., Homsa,F.,
Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsone,J., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,L.C.,
Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
Loulseghed,H., Lorzado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
Ma,J., Maheshwari,M., Mapua,P., Marondel,I., Martin,R.,
Martindale,A., Martinez,E., Massey,E., Mahoney,E., McLeod,M.P.,
Meador,M., Mei,G., Merscher,S., Metzker,M., Miller,A., Miner,G.,
Miner,Z., Mitchell,T., Mohabbat,K., Montgomery,K.T., Morgan,M.,
Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,N.,
Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
Ogulu,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shim,C.,
Shooshari,N., Sisson,I., Sodergren,E., Sonaike,T., Sparks,A.,
Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A.,
Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Teifrod,B.,
Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,
Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,
Washington,C., Watlington,S., Williams,G., Williams,A.,
Wlezcyk,R., Wooten,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,
Zorilla,S., Kucheriapati,R. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 145165)
Worley,K.C.
Direct Submission
Submitted (07-AUG-2000) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 145165)
Worley,K.C.
Direct Submission

JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Submitted (31-MAR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 145165)
Worley,K.C.
Direct Submission

Submitted (01-APR-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 145165)
Worley,K.C.
Direct Submission

Submitted (15-MAR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 145165)
Worley,K.C.
Direct Submission

Submitted (23-JAN-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Mar 31, 2001 this sequence version replaced gi:13324671.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URI:
<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

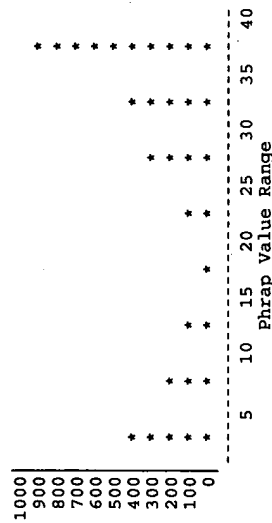
QUALSTAT-REPORT-----

----- Summary Statistics -----
Contig length: 145169
Phrap values in estimate: 144455
Average error rate (BCM-Phrap estimate): 5.28339e-05
Fraction of Phrap values less than 40 : 0.0137759
Number of consensus changing edits: 19
Number of N's in consensus : 0

----- Consensus changing edits -----
Position Original+Context Edited+Context

6882	agatgtggat (a) gctgtggctt	agatgtggat (g) gctgtggctt	QY	1695	GACCATAGCAGTTCTGGGACCTTACTTGTATGCTTTGGTGACTCAACAGATTTTCAGATC	1754
12828	ttcggctca (n) tgcacccctc	ttcggctca (c) tgcacccctc	Db	72478	GACCATAGCAGTTCTGGGACCTTACTTGTATGCTTTGGTGACTCAACAGATTTTCAGATC	72419
13151	aaacacacat (n) caatagataa	aaacacacat (a) caatagataa	QY	1755	CCCTCTGTACATTTTCCCCCATTAATAATGGGGTCATCGCAGCAGTGAACATGCT	1814
28957	ttgtattgca (n) atggcatctt	ttgtattgca (a) atggcatctt	Db	72418	CCCTCTGTACATTTTCCCCCATTAATAATGGGGTCATCGCAGCAGTGAACATGCT	72359
37139	atcgtattg (n) tgttaatttt	atcgtattg (t) tgttaatttt	QY	1815	CACGGAGTCTCCAGTGGCGTCTGGCTGGTGGATGGCTTCACTTTTGGCAGCTTTGC	1874
58030	tttttaatt (n) tatagttaat	tttttaatt (t) tatagttaat	Db	72358	CACGGAGTCTCCAGTGGCGTCTGGCTGGTGGATGGCTTCACTTTTGGCAGCTTTGC	72299
58290	tcattgaacta (n) tcattgacct	tcattgaacta (a) tcattgacct	QY	1875	ACGACATGGTGGCTGGTGGGAGAAATGGGGTGGTGGCCATGTCATTTTGGTCCAT	1934
64729	ctatctctac (n) aaagaaaga	ctatctctac (g) aaagaaaga	Db	72298	ACGACATGGTGGCTGGTGGGAGAAATGGGGTGGTGGCCATGTCATTTTGGTCCAT	72239
92472	aatggctttt (g) aaatgggccc	aatggctttt (a) aaatgggccc	QY	1935	TTTTGCTTCAGAAATCATCTGTTTCTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTC	1994
92662	cttagactgt (n) aatgcagga	cttagactgt (g) aatgcagga	Db	72238	TTTTGCTTCAGAAATCATCTGTTTCTCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTC	72179
103773	gctaattttt (n) catttttagt	gctaattttt (a) catttttagt	QY	1995	TGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTGAAAGTAATCAT	2054
104056	cacatgcatt (n) ggtatttgc	cacatgcatt (a) ggtatttgc	Db	72178	TGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTGAAAGTAATCAT	72119
117534	tcaagcttgt (n) atttcagggt	tcaagcttgt (a) atttcagggt	QY	2055	TTTGTCTGTGCTGCTGGCTTGAACATGGCCGAGTTCCTCTGCTGGTGGGAGCAA	2114
117581	cagatcacct (n) aggaatgct	cagatcacct (g) aggaatgct	Db	72118	TTTGTCTGTGCTGCTGGCTTGAACATGGCCGAGTTCCTCTGCTGGTGGGAGCAA	72059
125485	gcttacacta (n) aaatgaacc	gcttacacta (g) aaatgaacc	QY	2115	GTATGGCGCTCCCTCTCTGCTGCTTGGCTTTGGGGAGCCAGCACCATGGGCTA	2174
133723	agcatgggtg (n) nncacttaat	agcatgggtg (t) nncacttaat	Db	72058	GTATGGCGCTCCCTCTCTGCTGCTTGGCTTTGGGGAGCCAGCACCATGGGCTA	71999
133724	gcattggtgn (n) ncaacttaata	gcattggtgn (g) ncaacttaata	QY	2175	CATGCTGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATGACCATTCCTTACAC	2234
133725	catggtgtnn (n) cacttaataa	catggtgtnn (g) cacttaataa	Db	71998	CATGCTGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATGACCATTCCTTACAC	71939
140801	aaatattgtc (n) taaataaagt	aaatattgtc (c) taaataaagt	QY	2235	CAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATTTTGGGACTGCTCTATGCT	71879
			Db	71879	CAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATTTTGGGACTGCTCTATGCT	71879
			QY	2295	AAACACATTTGCCCTGCTTGTCTTCAACCACTGCATCTTAACTGCCCTTGTGGCTTCTT	2354
			Db	71878	AAACACATTTGCCCTGCTTGTCTTCAACCACTGCATCTTAACTGCCCTTGTGGCTTCTT	71819
			QY	2355	GTCTTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTGAAAGTAATTAAGTTATCCT	2414
			Db	71818	GTCTTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTGAAAGTAATTAAGTTATCCT	71759
			QY	2415	TCTGTGTAGTCCCACTTCTCTGCAATGCTCAATCCCTTCTCTACATCTTGTTCATGCC	2474
			Db	71758	TCTGTGTAGTCCCACTTCTCTGCAATGCTCAATCCCTTCTCTACATCTTGTTCATGCC	71699
			QY	2475	TCACCTTTAAGGAGGATCTGGTGGCTGAGAAAGCAAAACCTACGCTGGAGCAAGTCAA	2534
			Db	71698	TCACCTTTAAGGAGGATCTGGTGGCTGAGAAAGCAAAACCTACGCTGGAGCAAGTCAA	71639
			QY	2535	ACACCCAAAGCTTGATGTCAATTTAACTCTGATGATGTGCGAAACCAAGTCTCTGCACTCAAC	2594
			Db	71638	ACACCCAAAGCTTGATGTCAATTTAACTCTGATGATGTGCGAAACCAAGTCTCTGCACTCAAC	71579
			QY	2595	TCAAGCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTGCTCCAGTTCGGT	2654
			Db	71578	TCAAGCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTGCTCCAGTTCGGT	71519
			QY	2655	GCCATCAGCAGTTTATCCAGTGAAGCTGCTTCTTCTCTGCTGGGCAATTTGTCCC	2714
			Db	71518	GCCATCAGCAGTTTATCCAGTGAAGCTGCTTCTTCTCTGCTGGGCAATTTGTCCC	71459
			QY	2715	ATGCTCTTAA 2724	
			Db	71458	ATGCTCTTAA 71449	

----- Distribution of Quality < 40 Bases -----



Version: 1.01 qxf.
Location/Qualifiers
source

1. .145165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-186F10"

misc_feature

1. .2004
/note="overlaps bases 52259..54262 of clone AC073612"

/function="clone overlap"

/complement(6..223)

/rpt_family="LIME1"

/complement(404..698)

/rpt_family="LIM4"

/complement(833..1077)

/rpt_family="LIM4"

/complement(1228..1482)

/rpt_family="LIM4"

/complement(1740..2021)

/rpt_family="AluJb"

/complement(2098..2151)

/rpt_family="LIMEC"

/complement(2152..2449)

Query Match

Best Local Similarity 40.0%; Score 1090; DB 9; Length 145165;

Matches 1090; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1635 AGCCCTTCAACACCTGTGAACACCTGCTTGTGCTGGCTGATCAGAAATGGAGTGTG 1694

72538 AGCCCTTCAACACCTGTGAACACCTGCTTGTGCTGGCTGATCAGAAATGGAGTGTG 72479

QY

Db


```
PC C12N15/00,A61K37/02,C12N5/00
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..606
FT /organism='Unidentified'.
FEATURES
    source
        Location/Qualifiers
            1..606
            /organism="unidentified"
            /mol_type="genomic DNA"
            /db_xref="taxon:32644"
ORIGIN
Query Match 22.2%; Score 606; DB 6; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.9e-294;
Matches 606; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1684 ATTGGAGTGTGGACCATAGCAGTCTGGCAGCTTACTTGTAAATGCTTTGGTGAATCAACA 1743
Db 1 ATTGGAGTGTGGACCATAGCAGTCTGGCAGCTTACTTGTAAATGCTTTGGTGAATCAACA 60
QY 1744 GTTTTCAGATCCCTCTGTACATTTCCCCCATTAATGCTTTAAATGGGGTCAATCGCAGCA 1803
Db 61 GTTTTCAGATCCCTCTGTACATTTCCCCCATTAATGCTTTAAATGGGGTCAATCGCAGCA 120
QY 1804 GTGAACATGCTCACGGAGTCTCCAGTCCGCTGCTGGTGGTGGATCGGTTCACTTTT 1863
Db 121 GTGAACATGCTCACGGAGTCTCCAGTCCGCTGCTGGTGGTGGATCGGTTCACTTTT 180
QY 1864 GCGAGCTTTGACGACATGGTGGTGGGAGATGGGTTGGTGGCATGCTCATTTGGT 1923
Db 181 GCGAGCTTTGACGACATGGTGGTGGGAGATGGGTTGGTGGCATGCTCATTTGGT 240
QY 1924 TTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTGCGAGCCCTGGAG 1983
Db 241 TTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTGCGAGCCCTGGAG 300
QY 1984 CGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCCTG 2043
Db 301 CGTGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCCTG 360
QY 2044 AAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2103
Db 361 AAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
QY 2104 GGTGGCAGCAAGTATGGCCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2163
Db 421 GGTGGCAGCAAGTATGGCCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
QY 2164 ACCATGGGCTACATGGTGGCTCTCATTTGCTCAATTCCTTTGCTTCTCATGATGACC 2223
Db 481 ACCATGGGCTACATGGTGGCTCTCATTTGCTCAATTCCTTTGCTTCTCATGATGACC 540
QY 2224 ATTGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACTGGAGATATTTGGGAC 2283
Db 541 ATTGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACTGGAGATATTTGGGAC 600
QY 2284 TGCTCT 2289
Db 601 TGCTCT 606
```

Search completed: July 13, 2005, 00:42:34
Job time : 11604 secs

```
RESULT 15
AX016185
LOCUS AX016185 606 bp DNA linear PAT 07-SEP-2000
DEFINITION Sequence 3 from Patent EP0950711.
ACCESSION AX016185
VERSION AX016185.1 GI:10041794
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 14:17:43 ; Search time 1408 Seconds
(without alignments)
11452.681 Million cell updates/sec

Title: US-10-751-736-21

Perfect score: 2724
Sequence: 1 atggacacactccggctcgg.....cattgtcccatgtctctaa 2724

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2724	100.0	2724	3 Aaa30770	Aaa30770 Human G p
2	2724	100.0	2724	8 Abz42816	Abz42816 Human G p
3	2724	100.0	2724	10 Adc22782	Adc22782 Human G p
4	2724	100.0	2724	10 Adh14255	Adh14255 Human HG3
5	2724	100.0	2724	12 Adl32985	Adl32985 Human G p
6	2724	100.0	2724	12 Ado29923	Ado29923 Human GPC
7	2724	100.0	2880	11 Adn39796	Adn39796 Cancer/an
8	2724	100.0	2880	13 Adq80249	Adq80249 G protein
9	2724	100.0	2732	13 Adr67869	Adr67869 Human HG3
10	2724	100.0	3032	11 Adn39627	Adn39627 Cancer/an
11	2724	100.0	3032	11 Adn39530	Adn39530 Cancer/an
12	2724	100.0	4558	2 Aax23980	Aax23980 Human HG3
13	2724	100.0	4570	12 Adl12472	Adl12472 Human etc
14	2670	98.0	3438	10 Adf70582	Adf70582 Orphan re
15	2622	96.3	2724	3 Aaa30779	Aaa30779 DNA encod
16	2622	96.3	2724	10 Adc22796	Adc22796 Human G p
17	2622	96.3	2724	10 Adh14269	Adh14269 Mutated h
18	2603	95.6	3297	2 Aax23981	Aax23981 Human HG3
19	1864	68.4	2651	10 Adb80463	Adb80463 Ovarian c
20	1864	68.4	2651	11 Adn39165	Adn39165 Cancer/an

21	1864	68.4	2651	11	ADN39795	Adn39795 Cancer/an
22	1842	67.6	2082	2	AAZ25344	Aaz25344 Human LGR
23	699	25.7	723	8	ACA04772	ACA04772 cDNA enco
24	641	23.5	1790	6	ABQ5077	Abq5077 Human ova
25	606	22.2	606	2	Aaz40459	Aaz40459 Human gon
26	72	2.6	2636	5	ABA09697	Abas09697 Human bon
27	60	2.2	60	6	ABN47039	Abn47039 Human spl
28	41	1.5	2724	12	ADO30213	Ado30213 Mouse GPC
29	41	1.5	3098	13	ADR67870	Adr67870 Mouse HG3
30	30	1.1	30	12	ADI32988	Adi32988 5' FAM and
31	25	0.9	25	3	AAA30768	Aaa30768 Human G p
32	25	0.9	25	10	ADC22780	Adc22780 Human G p
33	25	0.9	25	10	ADH14253	Adh14253 Human HG3
34	24	0.9	30	3	AAA30767	Aaa30767 Human G p
35	24	0.9	30	10	ADC22779	Adc22779 Human G p
36	24	0.9	30	10	ADH14252	Adh14252 Human HG3
37	24	0.9	2901	6	AAI67921	Aai67921 Mouse LGR
38	24	0.9	2901	11	ADN02241	Adn02241 Mouse CDS
39	24	0.9	2901	12	ADK19406	Adk19406 ORF of CD
40	24	0.9	3637	6	AAI67920	Aai67920 Mouse LGR
41	24	0.9	3637	11	ADN02239	Adn02239 Mouse cDN
42	24	0.9	3637	12	ADK19404	Adk19404 cDNA enco
43	22	0.8	22	2	AAX23973	Aax23973 Human HG3
44	22	0.8	22	2	AAX23974	Aax23974 Human HG3
45	22	0.8	22	2	AAX23972	Aax23972 Human HG3

ALIGNMENTS

RESULT 1
AAA30770
ID AAA30770 standard; cDNA; 2724 BP.
AC AAA30770;
XX
XX
DT 21-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor HG38 cDNA.

KW G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening; agonist;
KW antagonist; ss.

OS Homo sapiens.

XX

XX WO200022129-A1.

XX PD 20-APR-2000.

XX PF 12-OCT-1999; 99WO-US023938.

XX PR 13-OCT-1998; 98US-00170496.

XX (AREN-) ARENA PHARM INC.

XX Behan DP, Chalmers DT, Liaw CW;

XX WPI; 2000-329165/28.

XX P-PSDB; AAY90682.

XX Non-endogenous constitutively activated human G protein-coupled
PT receptors, useful for identifying agonists for use as pharmaceutical
PT agents.
XX Example 1; Page 315-317; 341pp; English.

XX The invention relates to constitutively active, non-endogenous versions

CC of endogenous human orphan G protein-coupled receptors (GPCRs, AAY90643-
CC AAY90677 and AAY90683-Y90687) and to DNA encoding them (AAA30709-A30743
CC and AAA30775-A30779). The mutant proteins of the invention contain a
CC mutation in a portion of the protein comprising intracellular loop 3
CC (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, x,

QY	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
Db	661		720
QY	721	TTAAATACAAATAACCTTTGATGAATTCGCCCACTGCAATTAGGACACTCTCCAAACCTTTAAA	780
Db	721		780
QY	781	GAACTAGAGATTTATAGCAACAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC	840
Db	781		840
QY	841	CCTTCCTATTATACAAATACATTTCTATGACCAATCCATCCCAATTTGTTGGGAGATCTGCT	900
Db	841		900
QY	901	TTTCAACATTTACCTGAACTAAGAACTGACTGCTGAAATGGTGCCTCAAAAATACCTGAA	960
Db	901		960
QY	961	TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC	1020
Db	961		1020
QY	1021	TCATCTCTTCTCAAAACGCTCTGCAATCAGTTTACCTAATCTCCAAGTCTGATCTGTCT	1080
Db	1021		1080
QY	1081	TACAACTATTAGAAAGATTTTACCCAGTCTGCGCAAAAGCTTTAGAAATTTGAC	1140
Db	1081		1140
QY	1141	CTAAGACATTAATGAATCTAGCAAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC	1200
Db	1141		1200
QY	1201	CGATCGCTGAATTTGGCTTGGAAACAAATTCGTATTATTCACCCCAATGCAATTTTCCACT	1260
Db	1201		1260
QY	1261	TTGCCATCCCTAATAAGCTGGACCTATCGTCCAAACCTCTGTGCTCTTTTCCCTATAACT	1320
Db	1261		1320
QY	1321	GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATATCGCCCTTACAGAGCTTGATA	1380
Db	1321		1380
QY	1381	TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCGCTTATGCTTACCAGTGTGT	1440
Db	1381		1440
QY	1441	GCAATTTGGAGTGTGAGAAATGCTTAAAGTTTCTAATCAATGAATTAAGGTTGACAACT	1500
Db	1441		1500
QY	1501	AGCAGTATGGACACCTTCATAGAAAGATGCTGGAATGTTTCAAGCTCAAGATGAACGT	1560
Db	1501		1560
QY	1561	GACCTTGAAGATTTCTGCTTGCATTTGAGGAAGACCTGAAAGCCCTTCATTTCAGTGCAG	1620
Db	1561		1620
QY	1621	TGTTCACTTCCCAAGGCCCTTCAAACTCTGTGAAACCTGCTGATGCTGCTGATC	1680
Db	1621		1680
QY	1681	AGAAATGGAGTGTGACCAATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Db	1681		1740
QY	1741	ACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCAATTAACCTGTTAAATGGGGTCAATCGCA	1800
Db	1741		1800

Db	1741		1800
QY	1801	GCAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGCTGGTGTGTGAGTGCCTTCACT	1860
Db	1801		1860
QY	1861	TTTGGCAGCTTTGACCAACATGCTGGTGGGAGAAATGGGGTTGGTGCATGTCAAT	1920
Db	1861		1920
QY	1921	GGTTTTTGTGCTTTCAGAAATCATCTCTTTTCTGCTTACTCTGGCAGCCCTG	1980
Db	1921		1980
QY	1981	GAGCGTGGTCTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGC	2040
Db	1981		2040
QY	2041	CTGAAAGTAAATCATTTTGTCTGTGCGCTCTGCGCTTGAACATGGCGCGAGTTCCCTG	2100
Db	2041		2100
QY	2101	CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCGCTTGGCTTTTGGGGAGGCC	2160
Db	2101		2160
QY	2161	AGCAACATGGGCTACATGGTGCCTCATCTTGTCTAAATTCCTTTTGGTCTCATGATG	2220
Db	2161		2220
QY	2221	ACCAATGCCCTACCAAGCTCTAGTCAATTTTGACAAGGGAGACCTTGGAGAAATTTGG	2280
Db	2221		2280
QY	2281	GACTGCTCTATGTGTAACCAATTCATGCTCTTCCACCAACTGCATCCTTAAACTGC	2340
Db	2281		2340
QY	2341	CCTGTGGCTTTTGTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAGTA	2400
Db	2341		2400
QY	2401	ATTAAGTTTATCTCTTCTGGTGTAGTCCCACTCTCTCATGCTCTCAATCCCCTTCTTAC	2460
Db	2401		2460
QY	2461	ATCTTGTTCATCTCTCTTTAAGGAGATCTGGTGAAGCTGTGAGAGCAAACTTACGTC	2520
Db	2461		2520
QY	2521	TGGACAAGATCAAAACACCCAGCTTGTGATGTCATTAACCTGTGATGTCGAAAAACAG	2580
Db	2521		2580
QY	2581	TCCTGTGACTCAACTCAAGCTTGTAACTTACCAGCTCCAGCATCACATTTATGACCTG	2640
Db	2581		2640
QY	2641	CCTCCAGTTCCTGCGCCATCAACAGCTTATCCAGTGAAGCTGAGAGCTGCCATCTTCTCT	2700
Db	2641		2700
QY	2701	GTGGCATTTTGTCCCATGTCTCTAA	2724
Db	2701		2724

RESULT 4
ADH14255
ID ADH14255 standard; cDNA; 2724 BP.
XX
AC ADH14255;

Qy	361	CTGCAGAAATAAATCAGCTAAGACACGTTACCCACAGAAAGCTCTGCAGAAATTTTGCAGAGCCTT	420
Db	361	CTGCAGAAATAAATCAGCTAAGACACGTTACCCACAGAAAGCTCTGCAGAAATTTTGCAGAGCCTT	420
Qy	421	CAATCCCTCGGTCTGGATGCTTAAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
Db	421	CAATCCCTCGGTCTGGATGCTTAAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC	480
Qy	481	CTGCAATTCCTCTGAGGCACCTGTGTGCTGGATGACAATGCGTTTAAACAGAAATCCCGTCCAG	540
Db	481	CTGCAATTCCTCTGAGGCACCTGTGTGCTGGATGACAATGCGTTTAAACAGAAATCCCGTCCAG	540
Qy	541	GCTTTTAAAGTTTATCGGCAATTCGGAAGCCATGCTTGGCCCTGGAACAAAATACACCAAC	600
Db	541	GCTTTTAAAGTTTATCGGCAATTCGGAAGCCATGCTTGGCCCTGGAACAAAATACACCAAC	600
Qy	601	ATACCAGACTATGCTTTTGGAAACCTCTCCAGCTTTGGTAGTTCTACATCTCCCATTAACAAT	660
Db	601	ATACCAGACTATGCTTTTGGAAACCTCTCCAGCTTTGGTAGTTCTACATCTCCCATTAACAAT	660
Qy	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTTAGAT	720
Db	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTTAGAT	720
Qy	721	TTAAATTAACAATAACCTTTGATGAATTTCCCACTGSCAATTAGGACACTCTCCACCTTTAAA	780
Db	721	TTAAATTAACAATAACCTTTGATGAATTTCCCACTGSCAATTAGGACACTCTCCACCTTTAAA	780
Qy	781	GAACTAGGATTTCAATAGCAACAATATCAGGTGCGATAGCTGAGAAAGCAATTTGTAGGCAAC	840
Db	781	GAACTAGGATTTCAATAGCAACAATATCAGGTGCGATAGCTGAGAAAGCAATTTGTAGGCAAC	840
Qy	841	CTTCTCTTATTAACAATACTATTTCTATGACAAATCCCATCCAAATTTGTTGGAGACTGCT	900
Db	841	CTTCTCTTATTAACAATACTATTTCTATGACAAATCCCATCCAAATTTGTTGGAGACTGCT	900
Qy	901	TTTCAACAATTTTACCTGAACTAAGAAACACTGACTCTGAAATGGTGGCTCACAAAATAACTGAA	960
Db	901	TTTCAACAATTTTACCTGAACTAAGAAACACTGACTCTGAAATGGTGGCTCACAAAATAACTGAA	960
Qy	961	TTTCTGATTTTAACTGGAATCTGAAACCTGGAGAGTCTGACTTTTAACTGGAGCAGAGATC	1020
Db	961	TTTCTGATTTTAACTGGAATCTGAAACCTGGAGAGTCTGACTTTTAACTGGAGCAGAGATC	1020
Qy	1021	TCATCTCTTCTCTAAACCGGTCTGCAATCAGTTACTTAATCTCCAAAGTCTAGATCTGTCT	1080
Db	1021	TCATCTCTTCTCTAAACCGGTCTGCAATCAGTTACTTAATCTCCAAAGTCTAGATCTGTCT	1080
Qy	1081	TACAACCTATTAGAAGATTTTACCAGGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC	1140
Db	1081	TACAACCTATTAGAAGATTTTACCAGGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC	1140
Qy	1141	CTAAGACATAATGAAATCTACGAAATTAAGTTGACATTTTCCAGCAGAGTTGCTTAGCCTC	1200
Db	1141	CTAAGACATAATGAAATCTACGAAATTAAGTTGACATTTTCCAGCAGAGTTGCTTAGCCTC	1200
Qy	1201	CGATCGCTGAAATTTGGCTTGGACAAAATTTGCTATTATTACCCCAATGCAATTTCCACT	1260
Db	1201	CGATCGCTGAAATTTGGCTTGGACAAAATTTGCTATTATTACCCCAATGCAATTTCCACT	1260
Qy	1261	TTGGCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTGCTCTTTTCTATAACT	1320
Db	1261	TTGGCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTGCTCTTTTCTATAACT	1320
Qy	1321	GGGTTACATGGTTTAACTCACTTAAAAATTAAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Db	1321	GGGTTACATGGTTTAACTCACTTAAAAATTAAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Qy	1381	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCGTTATGCTTACCAAGTGTCT	1440
Db	1381	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCGTTATGCTTACCAAGTGTCT	1440

QY	1441	GCATTTGGAGTGTGTGAGAAATGCCTATAAGATTTCTAATCAATGGAATTAAGGTGACAAC	1500
DB	1441	GCATTTGGAGTGTGTGAGAAATGCCTATAAGATTTCTAATCAATGGAATTAAGGTGACAAC	1500
QY	1501	AGCAGTATGGACGACCTTCATAGAAGATGCTGGAAATGTTTCAGAGCTCAAGATGAACGT	1560
DB	1501	AGCAGTATGGACGACCTTCATAGAAGATGCTGGAAATGTTTCAGAGCTCAAGATGAACGT	1560
QY	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTCGAAAGCCCTTCATTCAAGTGCAG	1620
DB	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTCGAAAGCCCTTCATTCAAGTGCAG	1620
QY	1621	TGTTCACTTCCGAGGCCCTTCAAACCTGTGGAACACCTGCTTGATGGCTGGCTGATC	1680
DB	1621	TGTTCACTTCCGAGGCCCTTCAAACCTGTGGAACACCTGCTTGATGGCTGGCTGATC	1680
QY	1681	AGAAATGGAGTGTGGACCATAGCAGTCTTCGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
DB	1681	AGAAATGGAGTGTGGACCATAGCAGTCTTCGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
QY	1741	ACAGTTTTTCAGATCCCCCTCTGTACATTTCCGCCCACTAAACCTGTTAAATGGGGTCAATCGCA	1800
DB	1741	ACAGTTTTTCAGATCCCCCTCTGTACATTTCCGCCCACTAAACCTGTTAAATGGGGTCAATCGCA	1800
QY	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGAGATGGGGTGGTGGCCATGTCACT	1860
DB	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGCTGGCTGGAGATGGGGTGGTGGCCATGTCACT	1860
QY	1861	TTTGGCAGCTTTGCAACACATGCTGCTGGTGGAGAAATGGGGTGGTGGCCATGTCACT	1920
DB	1861	TTTGGCAGCTTTGCAACACATGCTGCTGGTGGAGAAATGGGGTGGTGGCCATGTCACT	1920
QY	1921	GGTTTTTTTGTCCATTTTTTGTCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGGCCCTG	1980
DB	1921	GGTTTTTTTGTCCATTTTTTGTCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGGCCCTG	1980
QY	1981	GAGCGTGGGTCTCTGTGAAATAATCTCGAAAAATTTGAAAAAGCTCCATTTTCTAGC	2040
DB	1981	GAGCGTGGGTCTCTGTGAAATAATCTCGAAAAATTTGAAAAAGCTCCATTTTCTAGC	2040
QY	2041	CTGAAAGTAATCATTTTGCTCTGTGCCCTGCTGGCCTTGACCAATGGCCGAGTTCCTCCCTG	2100
DB	2041	CTGAAAGTAATCATTTTGCTCTGTGCCCTGCTGGCCTTGACCAATGGCCGAGTTCCTCCCTG	2100
QY	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGTGCTTTTGGCTTTGGGGAGCCC	2160
DB	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGTGCTTTTGGCTTTGGGGAGCCC	2160
QY	2161	AGCACCATGGGTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTTGCTTTCTCTCATGATG	2220
DB	2161	AGCACCATGGGTACATGGTGGCTCTCATCTTGTCTCAATTTCCCTTTGCTTTCTCTCATGATG	2220
QY	2221	ACCAATGGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTTGG	2280
DB	2221	ACCAATGGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTTGG	2280
QY	2281	GACTGCTCTATGTTAAACAACATTTGCCCTGCTCTTTCACCAACTGCACTCTTAAACTGC	2340
DB	2281	GACTGCTCTATGTTAAACAACATTTGCCCTGCTCTTTCACCAACTGCACTCTTAAACTGC	2340
QY	2341	CCTGTGGCTTTCTTGCTCTCTCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA	2400
DB	2341	CCTGTGGCTTTCTTGCTCTCTCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA	2400
QY	2401	ATTAAGTTTTATCTTCTGTGGTGTAGTCCCACTCTTCGCAATGCTCAATCCCTTCTCTAC	2460
DB	2401	ATTAAGTTTTATCTTCTGTGGTGTAGTCCCACTCTTCGCAATGCTCAATCCCTTCTCTAC	2460
QY	2461	ATCTTGTTCATCTCTCATCTTTAAGGAGGATCTGGTGAGCTGGAAGCAAACTTACGTC	2520
DB	2461	ATCTTGTTCATCTCTCATCTTTAAGGAGGATCTGGTGAGCTGGAAGCAAACTTACGTC	2520
QY	2521	TGGAACAAGATCAAAACACCCCAAGCTTGATGCTCAATTAAGCTCTGATGATGCGAAAAACAG	2580

Qy	721	TTAAATTACAATAAACCCTTGATGAATTTCCCACTGGAATTAGBACACTCTCCAACTTTAAA	780
Db	769	TTAAATTACAATAAACCCTTGATGAATTTCCCACTGGAATTAGBACACTCTCCAACTTTAAA	828
Qy	781	GAACTAGAGTTTCATAGCAACAATAATCAGGTGCGATCTGAGTAAGACATTTGTAGGCAAC	840
Db	829	GAACTAGAGTTTCATAGCAACAATAATCAGGTGCGATCTGAGTAAGACATTTGTAGGCAAC	888
Qy	841	CCCTCTCTTAATTACAATAACATTTCTATGACAAATCCCATCCAAATTTGTGGAGATCTGCT	900
Db	889	CCCTCTCTTAATTACAATAACATTTCTATGACAAATCCCATCCAAATTTGTGGAGATCTGCT	948
Qy	901	TTTTCACACATTTTACCTGGAACCTAAGAACACACTGACTCTGAAATGGTGCCTCACAAATACAGAA	960
Db	949	TTTTCACACATTTTACCTGGAACCTAAGAACACACTGACTCTGAAATGGTGCCTCACAAATACAGAA	1008
Qy	961	TTTTCCTGATTTAACTGGAACTGCAAACTGAGAGAGTCTGACTTTTAACTGAGGACACAGATC	1020
Db	1009	TTTTCCTGATTTAACTGGAACTGCAAACTGAGAGAGTCTGACTTTTAACTGAGGACACAGATC	1068
Qy	1021	TCATCTCTTCTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCAAGTGTCTAGATCTGTCT	1080
Db	1069	TCATCTCTTCTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCAAGTGTCTAGATCTGTCT	1128
Qy	1081	TAGAACCTATTAGAAAGATTTTACCACAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGCAC	1140
Db	1129	TAGAACCTATTAGAAAGATTTTACCACAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGCAC	1188
Qy	1141	CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACACTTTCACAGCAGTTGCTTAGGCCTC	1200
Db	1189	CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACACTTTCACAGCAGTTGCTTAGGCCTC	1248
Qy	1201	CGATCGCTGAATTTGGCTTGGAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACT	1260
Db	1249	CGATCGCTGAATTTGGCTTGGAACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACT	1308
Qy	1261	TTGCGCATCCCTATATAAGCTGGACCTATCGTCCAACTCTGTGCTCTTTTCTCTATAACT	1320
Db	1309	TTGCGCATCCCTATATAAGCTGGACCTATCGTCCAACTCTGTGCTCTTTTCTCTATAACT	1368
Qy	1321	GGGTTACATGGTTTAACTCACTTAAAAATTAAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Db	1369	GGGTTACATGGTTTAACTCACTTAAAAATTAAACAGGAAATCATGCTTTACAGAGCTTGATA	1428
Qy	1381	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACAGTGTCTGT	1440
Db	1429	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACAGTGTCTGT	1488
Qy	1441	GCATTTGGAGTGTGTAGAGATGCTTATAAGATTTCTAATCAATGGAATTAAGGTGACAAAC	1500
Db	1489	GCATTTGGAGTGTGTAGAGATGCTTATAAGATTTCTAATCAATGGAATTAAGGTGACAAAC	1548
Qy	1501	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1560
Db	1549	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT	1608
Qy	1561	GACCTTGAAGATTTCTGTCTTGACTTTGAGGAGAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
Db	1609	GACCTTGAAGATTTCTGTCTTGACTTTGAGGAGAGACCTGAAAGCCCTTCAATCAGTGCAG	1668
Qy	1621	TGTTTCACTTCCCGACGCCCTTCAAAACCTGTGGAACACCTGCTTGATGCTCGCTGTATC	1680
Db	1669	TGTTTCACTTCCCGACGCCCTTCAAAACCTGTGGAACACCTGCTTGATGCTCGCTGTATC	1728
Qy	1681	AGAAATGGAGTGTGGACCATAGCAGTTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Db	1729	AGAAATGGAGTGTGGACCATAGCAGTTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1788
Qy	1741	ACAGTTTTTCAGATCCCTCTGTGATCATTTTCCCGCAATTAACATGTTTAATTTGGGGTCAATCGCA	1800
Db	1789	ACAGTTTTTCAGATCCCTCTGTGATCATTTTCCCGCAATTAACATGTTTAATTTGGGGTCAATCGCA	1848

DEC 11 1966

RESULT 8
AD090340

ADQ80249
ID ADQ80249 standard; cDNA; 2880 BP.

2000

AC ADQ80249;

XX

13

1441 GCATTTGGAGTGTGAGAAATGCTATAAGATTTCTAAATCAATGGAATAAAGTGACAAC 1500
1489 GCATTTGGAGTGTGAGAAATGCTATAAGATTTCTAAATCAATGGAATAAAGTGACAAC 1548
1501 AGCAGTATGACCACTTCATAGAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
1549 AGCAGTATGACCACTTCATAGAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1608
1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGACG 1620
1609 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAAGTGACG 1668
1621 TGTTCACCTTCCCGAGCCCTTCAAACCCCTGTAACACCTGTAACACCTGTAACACCTGTAAC 1680
1669 TGTTCACCTTCCCGAGCCCTTCAAACCCCTGTAACACCTGTAACACCTGTAACACCTGTAAC 1728
1681 AGAATTTGGAGTGTGGAACCATAGCAGTTCTGGACATTTCTGTAATGCTTTGGTGAATCA 1740
1729 AGAATTTGGAGTGTGGAACCATAGCAGTTCTGGACATTTCTGTAATGCTTTGGTGAATCA 1788
1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAATGCGGGTCATCGCA 1800
1789 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAATGCGGGTCATCGCA 1848
1801 GCAGTGAAATGCTACAGGAGTCTCCAGTGGCGTCTGGCTGGTGGTGGTGGTGGTGGTGGTGG 1860
1849 GCAGTGAAATGCTACAGGAGTCTCCAGTGGCGTCTGGCTGGTGGTGGTGGTGGTGGTGGTGG 1908
1861 TTTGGCAGCTTTGCAGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGTGG 1920
1909 TTTGGCAGCTTTGCAGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGTGG 1968
1921 GGTGTTTTGTCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 1980
1969 GGTGTTTTGTCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 2028
1981 GAGCGTGGTGTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
2029 GAGCGTGGTGTCTGTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2088
2041 CTGAAAGTATCATTTTGTCTGTGCGCTGCTGGCTTGAACATGGCCGAGTCCCTG 2100
2089 CTGAAAGTATCATTTTGTCTGTGCGCTGCTGGCTTGAACATGGCCGAGTCCCTG 2148
2101 CTGGGTGGCAGCAAGTATGCGCTCCCTCTCTGCTGCTGCTTGTGCTTGTGGGAGCC 2160
2149 CTGGGTGGCAGCAAGTATGCGCTCCCTCTCTGCTGCTGCTTGTGCTTGTGGGAGCC 2208
2161 AGCACCATGGCTTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2220
2209 AGCACCATGGCTTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2268
2221 ACATTTGCTTACACCAAGCTTCTGCAATTTGGACAGGAGACCTGGAGAAATTTGG 2280
2269 ACCATTTGCTTACACCAAGCTTCTGCAATTTGGACAGGAGACCTGGAGAAATTTGG 2328
2281 GACTGCTCTATGTAACACATGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 2340
2329 GACTGCTCTATGTAACACATGCTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 2388
2341 CTGTGGCTTCTTGTGCTTCTCTCTTAAATAAACCTTTACATTTATCATGCTGCAAGTA 2400
2389 CCGTGGCTTCTTGTGCTTCTCTCTTAAATAAACCTTTACATTTATCATGCTGCAAGTA 2448
2401 ATTAAGTTTATCTTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
2449 ATTAAGTTTATCTTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2508
2461 ATCTGTTCAATCTCACTTTAAGGAGGATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2520
2509 ATCTGTTCAATCTCACTTTAAGGAGGATCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2568

2521 TGGACAGATCAAAACACCAAGCTTGATGTCATTAATCTGATGATGTCGAAAAACAG 2580
2569 TGGACAGATCAAAACACCAAGCTTGATGTCATTAATCTGATGATGTCGAAAAACAG 2628
2581 TCCTGTGACTCAACTCAAGCCCTTGTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2640
2629 TCCTGTGACTCAACTCAAGCCCTTGTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2688
2641 CTTCCAGTTCGTCGCCATCACAGCTTATCCAGTGAAGCTGAGAGCTGCCATCTTCTCTCT 2700
2689 CTTCCAGTTCGTCGCCATCACAGCTTATCCAGTGAAGCTGAGAGCTGCCATCTTCTCTCT 2748
2701 GTGGCATTTGTCCTCATGTCCTAA 2724
2749 GTGGCATTTGTCCTCATGTCCTAA 2772

RESULT 9
ADR67869
ID ADR67869 standard; cDNA; 2973 BP.
XX
AC ADR67869;
XX
DT 18-NOV-2004 (first entry)
XX
DE Human HG38 coding sequence.
XX
KW ss: gene; human ; G protein-coupled receptor ; GPCR; HG38; colon ; lung ;
cancer.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 18..2741
FT /*tag= a
FT /product= "Human HG38"
XX
PN WO2004074436-A2.
XX
PD 02-SEP-2004.
XX
PF 11-FEB-2004; 2004WO-US004060.
XX
PR 19-FEB-2003; 2003US-0448959P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Lasek AW;
XX
DR WPI; 2004-652946/63.
DR P-PSDB; ADR67868.
XX
PT Detecting colon or lung cancer, by determining amount of protein in
sample, comparing amount of protein to standard, and differential
expression of protein in sample indicates colon or lung cancer.
PT
XX
PS Claim 1; SEQ ID NO 2; 79pp; English.
XX
CC This sequence encodes the human G protein-coupled receptor (GPCR) known
as HG38. The HG38 protein and corresponding nucleic acid, may be used in
the method of the invention for detecting colon or lung cancer. The first
method involves performing an assay to determine the amount of HG38 in a
sample of colon or lung tissue, and comparing the amount of protein to
standard, thus detecting expression of protein in sample, where
differential expression of protein in sample when compared with the
standard is diagnostic of colon or lung cancer. The second method
involves hybridizing a composition comprising the HG38 coding sequence,
or its complement, and a labelling moiety, to nucleic acids of a sample
of colon or lung tissue under conditions to form at least one
hybridization complex, detecting hybridization complex formation, and
comparing complex formation to a standard, where the comparison reflects
differential expression of the polynucleotide in the sample relative to
the standard and is diagnostic of a colon or lung cancer. This first

Db	2058		CTGAAAGTAATCATTTTGCTCTGTGCCTGTGCCTTGCACCATGGCGCAGTAGTTCCCTCG	2117
Qy	2101		CTGGTGGCAGCAAGTATGGCGCCTCCCTCTCTGCCTGCCTTTGGCTTTTGGGAGCCC	2160
Db	2118		CTGGTGGCAGCAAGTATGGCGCCTCCCTCTCTGCCTGCCTTTGGGAGCCC	2177
Qy	2161		AGCACCATGGGCTACATGGTCGCTCTCATCTTGTCTCAATCCCTTTGCTTCCTCATGATG	2220
Db	2178		AGCACCATGGGCTACATGGTCGCTCTCATCTTGTCTCAATCCCTTTGCTTCCTCATGATG	2237
Qy	2221		ACCATGGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGATATTTGG	2280
Db	2238		ACCATGGCTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGATATTTGG	2297
Qy	2281		GACTGCTCTATGGTAAACAATTTGCCCTGTGTCTTTCACCAACTGCATCCTAAACTGC	2340
Db	2298		GACTGCTCTATGGTAAACAATTTGCCCTGTGTCTTTCACCAACTGCATCCTAAACTGC	2357
Qy	2341		CCTGTGGCTTTCTGTGCTTCTCTCTTTTAAATAAACCTTACAATTAATCAGTCTCGAAGTA	2400
Db	2358		CCTGTGGCTTTCTGTGCTTCTCTCTTTTAAATAAACCTTACAATTAATCAGTCTCGAAGTA	2417
Qy	2401		ATTTAAGTTTATCCTTCTGGTGGTAGTCCCACTTCCCTGCATGTCTCAATCCCTTCTCTAC	2460
Db	2418		ATTTAAGTTTATCCTTCTGGTGGTAGTCCCACTTCCCTGCATGTCTCAATCCCTTCTCTAC	2477
Qy	2461		ATCTTGTTCATCTCTCACTTTAAGCAGGATCTGTGTGAGCCTTGAGAAAGCAAACTACGTC	2520
Db	2478		ATCTTGTTCATCTCTCACTTTAAGCAGGATCTGTGTGAGCCTTGAGAAAGCAAACTACGTC	2537
Qy	2521		TGGAACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTGTGATGATGTCGAAAACAG	2580
Db	2538		TGGAACAAGATCAAAACACCCCAAGCTTGATGTCAATTAACCTGTGATGATGTCGAAAACAG	2597
Qy	2581		TCCTGTGACTCAACTCAAGCCTTGGTAACTTTACAGCTCCAGCATCACTTATGACCTG	2640
Db	2598		TCCTGTGACTCAACTCAAGCCTTGGTAACTTTACAGCTCCAGCATCACTTATGACCTG	2657
Qy	2641		CCTCCCAAGTCCGTCGCATCAACAGCTTTATCCAGTGACTGAGAGCTGCCATCTTTCTCT	2700
Db	2658		CCTCCCAAGTCCGTCGCATCAACAGCTTTATCCAGTGACTGAGAGCTGCCATCTTTCTCT	2717
Qy	2701		GTGGCATTTGTGCCATGTCTCTAA	2724
Db	2718		GTGGCATTTGTGCCATGTCTCTAA	2741

RESULT 10

RESULTS IN
ADN39627
ID ADN39627 standard; cDNA; 3032 BP.

XX
AC

17-JUN-2004 (first entry)

XX
DE Cancer/angiogenesis/fibrosis-related nucleic acid. SEO ID NO:A227.

Human; differential expression; cancer; angiogenic disorder; fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis; inflammatory disease; autoimmune disease; retinal neovascularisation syndrome; scarring; uterine fibroid; detection; diagnosis; prognosis; drug screening; drug targeting; wound healing; contraception; cytostatic; cardiac; immunomodulatory; vulvular; gene therapy; vaccine; gene; ss.

XX
OS
Homo sapiens.

XX PN WO2003042661-A2.

XX
PD 22-MAY-2003

13-NOV-2002: 2002WO-US036810.

XX	13-NOV-2001;	2001IUS-0350666P.
PR	21-NOV-2001;	2001IUS-0332464P.
PR	29-NOV-2001;	2001IUS-0334393P.
PR	03-DEC-2001;	2001IUS-0335394P.
PR	14-DEC-2001;	2001IUS-0340376P.
PR	08-JAN-2002;	2001IUS-0347211P.
PR	10-JAN-2002;	2002IUS-0347349P.
PR	18-FEB-2002;	2002IUS-035250P.
PR	03-FEB-2002;	2002IUS-0356714P.
PR	20-FEB-2002;	2002IUS-0359077P.
PR	29-MAR-2002;	2002IUS-0368809P.
PR	04-APR-2002;	2002IUS-0370110P.
PR	12-APR-2002;	2002IUS-0372246P.
PR	03-JUN-2002;	2002IUS-0386614P.
PR	16-JUL-2002;	2002IUS-0396839P.
PR	22-JUL-2002;	2002IUS-0397775P.
PR	29-AUG-2002;	2002IUS-0397845P.
PR	09-SEP-2002;	2002IUS-0409450P.

(EOSB-) EOS BIOTECHNOLOGY INC.

Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA; Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

WPI; 2003-468649/44.

Determining the presence or absence of a pathological cell in a patient, useful for diagnosing, prognosing or treating cancer, comprises detecting a nucleic acid in a biological sample.

Claim 8: SEO ID NO A227: 1385pp: English.

The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers and other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularisation syndromes, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a nucleic acid sequence of the invention.

Sequence 3032 BP: 729 A: 816 C: 648 G: 839 T: 0 U: 0 Other: XX

Query Match	100.0%;	Score 2724;	DB 11;	Length 3032;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2724:	Conservative	0;	Mismatches	0;
	Indels	0;	Gaps	0;

Qy	1	ATG	GAC	CTC	CGG	CTC	GTG	TCT	CTC	TGT	CGC	TGT	GTG	TCT	GCA	GTC	GAC	TGG	CGA	CC	60
Dp	201	ATG <td>GAC <td>CTC <td>CGG <td>CTC <td>GTG <td>TCT <td>CTC <td>TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	GAC <td>CTC <td>CGG <td>CTC <td>GTG <td>TCT <td>CTC <td>TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CTC <td>CGG <td>CTC <td>GTG <td>TCT <td>CTC <td>TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CGG <td>CTC <td>GTG <td>TCT <td>CTC <td>TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	CTC <td>GTG <td>TCT <td>CTC <td>TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	GTG <td>TCT <td>CTC <td>TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td></td></td></td>	TCT <td>CTC <td>TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td></td></td>	CTC <td>TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td></td>	TGT <td>CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td></td>	CGC <td>TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td></td>	TGT <td>GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td></td>	GTG <td>TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td></td>	TCT <td>GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td></td>	GCA <td>GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td></td>	GTC <td>GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td></td>	GAC <td>TGG <td>CGA <td>CC</td> <td>260</td> </td></td>	TGG <td>CGA <td>CC</td> <td>260</td> </td>	CGA <td>CC</td> <td>260</td>	CC	260

Qy	61	GGGGGAGCTCTCCAGTCTGGTGTGTGCTGAGGGGTGCCCCACACACTGTCATTGC	120
Dp	261	GGGGGAGCTCTCCAGTCTGGTGTGTGCTGAGGGGTGCCCCACACACTGTCATTGC	320

Qy	121	GAGCCCGACGGCAGGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGAGCTG	180
Dh	321	GAGCCCGACGGCAGGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGAGCTG	380

Qy	181	CTTTCACCTCAGCGCTTTCACCTCTTACCTAGACCTCAGTATGAACAA	CATCAGTCAG	240
Ph	281	CTTTCGACCTCAGCGCTTTCACCTCTTACCTAGACCTCAGTATGAACAA	CATCAGTCAG	440

QY 241 CTGCTCCGGAATCCCTCGCCAGTCTCCGCTTCCTGGAGGATTAAGTCTTGGCGGAAC 300
Db 441 CTGCTCCGGAATCCCTCGCCAGTCTCCGCTTCCTGGAGGATTAAGTCTTGGCGGAAC 500
QY 301 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCCTTACAGTCTTAAAGTCTTATG 360
Db 501 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCCTTACAGTCTTAAAGTCTTATG 560
QY 361 CTGCAGAAATAATCAGCTAAGACACAGTACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db 561 CTGCAGAAATAATCAGCTAAGACACAGTACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 620
QY 421 CAATCCCTGCGTCTGGATGCTTAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 480
Db 621 CAATCCCTGCGTCTGGATGCTTAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 680
QY 481 CTGCATTCCTTGAGGACCTGTGGCTGGATGACATGGTTAAACAGAAATCCCGTCCAG 540
Db 681 CTGCATTCCTTGAGGACCTGTGGCTGGATGACATGGTTAAACAGAAATCCCGTCCAG 740
QY 541 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
Db 741 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 800
QY 601 ATACAGAGATATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCGATACAAAT 660
Db 801 ATACAGAGATATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCGATACAAAT 860
QY 661 AGAATCCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCTTAGAGACTTTAGAT 720
Db 861 AGAATCCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCTTAGAGACTTTAGAT 920
QY 721 TTAATTTACAATAAATCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
Db 921 TTAATTTACAATAAATCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 980
QY 781 GAACTAGGATTTATAGCAACAAATATCAGTGCATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 981 GAACTAGGATTTATAGCAACAAATATCAGTGCATACCTGAGAAAGCAATTTGTAGGCAAC 1040
QY 841 CCTCTCTTATTAACAATATTTCTATGACAAATCCCACTTAAATTTGTTGGAGATCTGCT 900
Db 1041 CCTCTCTTATTAACAATATTTCTATGACAAATCCCACTTAAATTTGTTGGAGATCTGCT 1100
QY 901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGCCTCAAAATAACTGAA 960
Db 1101 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGCCTCAAAATAACTGAA 1160
QY 961 TTTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 1161 TTTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1220
QY 1021 TCATCTCTTCCCTCAAACTGCTGCAATCAGTTTACCTTAATCTCAAGTGTAGATCTGCT 1080
Db 1221 TCATCTCTTCCCTCAAACTGCTGCAATCAGTTTACCTTAATCTCAAGTGTAGATCTGCT 1280
QY 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATGAC 1140
Db 1281 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATGAC 1340
QY 1141 CTAAGACATAATGAATCTACGAATAATGAATGACATTTTCCAGCAGTGTGCTTAGCCTC 1200
Db 1341 CTAAGACATAATGAATCTACGAATAATGAATGACATTTTCCAGCAGTGTGCTTAGCCTC 1400
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTTACCCCAATGCAATTTCCACT 1260
Db 1401 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTTACCCCAATGCAATTTCCACT 1460
QY 1261 TTGCCATCCCTAATAAAGCTGAGCCTATCCGTCAACCTCCTGCTGCTTTTCCCTAATACT 1320
Db 1461 TTGCCATCCCTAATAAAGCTGAGCCTATCCGTCAACCTCCTGCTGCTTTTCCCTAATACT 1520

QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCAATGCTTACAGAGCTTGATA 1380
Db 1521 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCAATGCTTACAGAGCTTGATA 1580
QY 1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTGT 1440
Db 1581 TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGTGT 1640
QY 1441 GCATTTGAGTGTGTGAGAAATGCTTATAGAAATTTCTAATCAATGGAATAAAGTGACAAAC 1500
Db 1641 GCATTTGAGTGTGTGAGAAATGCTTATAGAAATTTCTAATCAATGGAATAAAGTGACAAAC 1700
QY 1501 AGCAGTATGAGACGACCTTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db 1701 AGCAGTATGAGACGACCTTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1760
QY 1561 GACCTTGAAGATTTTCTGCTTGAATTTTTCAGGAAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1761 GACCTTGAAGATTTTCTGCTTGAATTTTTCAGGAAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1820
QY 1621 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGCTGTGATGGCTGCTGATC 1680
Db 1821 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGCTGTGATGGCTGCTGATC 1880
QY 1681 AGAATGAGTGTGAGACCATAGCAGTCTGAGCACTTACTTCTAATGCTTTCGTGACTTCA 1740
Db 1881 AGAATGAGTGTGAGACCATAGCAGTCTGAGCACTTACTTCTAATGCTTTCGTGACTTCA 1940
QY 1741 ACAGTTTTTCAGATCCCTCTGTATCAATTTCCCAATTTAAACTGTTAAATTTGGGGTCAATGCA 1800
Db 1941 ACAGTTTTTCAGATCCCTCTGTATCAATTTCCCAATTTAAACTGTTAAATTTGGGGTCAATGCA 2000
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGGCTGGTGGTGGATCGGTTCACT 1860
Db 2001 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGGCTGGTGGTGGATCGGTTCACT 2060
QY 1861 TTTGCGAGCTTTGACACGACATGCTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAAT 1920
Db 2061 TTTGCGAGCTTTGACACGACATGCTGCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCAAT 2120
QY 1921 GGTTTTTTGTCCATTTTTTGTCTCAGAAATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db 2121 GGTTTTTTGTCCATTTTTTGTCTCAGAAATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 2180
QY 1981 GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCCAATTTCTAGC 2040
Db 2181 GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCCAATTTCTAGC 2240
QY 2041 CTGAAAGTAAATCATTTTTTGTCTCTGCTGCTGCTGGCCTTGACCATGGCCGACAGTTCCCTG 2100
Db 2241 CTGAAAGTAAATCATTTTTTGTCTCTGCTGCTGCTGGCCTTGACCATGGCCGACAGTTCCCTG 2300
QY 2101 CTGGGTGGCAGCAAGATATGGCCCTCCCTCTCTGCTGCTGCTTGGCCCTTTTGGGGAGGCC 2160
Db 2301 CTGGGTGGCAGCAAGATATGGCCCTCCCTCTCTGCTGCTGCTTGGCCCTTTTGGGGAGGCC 2360
QY 2161 AGCACCATTGGCTACATGGTCTCATCTTGTCTCAATTTCCCTTGTCTTCTCATGATG 2220
Db 2361 AGCACCATTGGCTACATGGTCTCATCTTGTCTCAATTTCCCTTGTCTTCTCATGATG 2420
QY 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATTTGG 2280
Db 2421 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATTTGG 2480
QY 2281 GACTGCTCTATGGTAAAAACAATGGCCCTGTTGCTCTTCAACCACTGATCTCTAACTGC 2340
Db 2481 GACTGCTCTATGGTAAAAACAATGGCCCTGTTGCTCTTCAACCACTGATCTCTAACTGC 2540
QY 2341 CTTGTGGCTTTCTTGTGCTCTCTCTTAAATTAACCTTACATTTATCAGTCTCTCAAGTA 2400
Db 2541 CTTGTGGCTTTCTTGTGCTCTCTCTTAAATTAACCTTACATTTATCAGTCTCTCAAGTA 2600
QY 2401 ATTAAGTTTATCTCTTCTGCTGGTAGTCCCACTTCTCTGATGCTCTCAATCCCTTCTCTAC 2460

Query Match	Best Local Similarity	100.0%;	Score 2724;	DB 12;	Length 4570;
Matches 2724;	Conservative	0;	Mismatches	0;	Gaps 0;
QY	1	ATGGACACCTCCCGGCTCGGTGCTCCTGTCCTTGCCTGCTGCTGCTGCTGCGAGCC	60		
DB	262	ATGGACACCTCCCGGCTCGGTGCTCCTGTCCTTTGCTGCTGCTGCTGCTGCGAGCC	321		
QY	61	GGGGGAGCTCTCCCAGGCTCTGGTGTTGCTGAGGGGCTGCCCAACAACACTGTCATTGC	120		
DB	322	GGGGGAGCTCTCCCAGGCTCTGGTGTTGCTGAGGGGCTGCCCAACAACACTGTCATTGC	381		
QY	121	GAGCCGACGGCAGGATGTTGCTCAGGGTGAAGTGCTCCGACTGGGGCTCTCGGAGCTG	180		
DB	382	GAGCCGACGGCAGGATGTTGCTCAGGGTGAAGTGCTCCGACTGGGGCTCTCGGAGCTG	441		
QY	181	CCTTCCAACCTCAGGCTCTCACCTCTACCTAGACCTCAGTATGAACAACATCAGTCAG	240		
DB	442	CCTTCCAACCTCAGGCTCTCACCTCTACCTAGACCTCAGTATGAACAACATCAGTCAG	501		
QY	241	CTGCTCCCCAAATCCCGCCAGCTCCCGCTTCCTCGAGGAGTTACGTTCTTGGGGAAAC	300		
DB	502	CTGCTCCCCAAATCCCGCCAGCTCCCGCTTCCTCGAGGAGTTACGTTCTTGGGGAAAC	561		
QY	301	GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGSCCTTTACAGTCTTTAAAGTTCTTATG	360		
DB	562	GCTCTGACATACATTTCCCAAGGGAGCATTCACCTGSCCTTTACAGTCTTTAAAGTTCTTATG	621		
QY	361	CTGCAGAATAAATCAGCTTAAGACACGTAACCACAGAGCTCTGAGAAATTTGCGAAGCTT	420		
DB	622	CTGCAGAATAAATCAGCTTAAGACACGTAACCACAGAGCTCTGAGAAATTTGCGAAGCTT	681		
QY	421	CAATCCCTCGCTGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAGTGCC	480		
DB	682	CAATCCCTCGCTGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAGTGCC	741		
QY	481	CTGCATTTCCCTGAGGCACCTGTTGGCTGGAAGACAAATGCGTTAACAGAAATCCCGCTCCAG	540		
DB	742	CTGCATTTCCCTGAGGCACCTGTTGGCTGGAAGACAAATGCGTTAACAGAAATCCCGCTCCAG	801		
QY	541	GCTTTTTAGAAATTTAGGCAATGCAAGCAATGACCTTGGGCTTCCAGAAATCCCGCTCCAG	600		
DB	802	GCTTTTTAGAAATTTAGGCAATGCAAGCAATGACCTTGGGCTTCCAGAAATCCCGCTCCAG	861		
QY	601	ATACACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTCCATAACAAT	660		
DB	862	ATACACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTCCATAACAAT	921		
QY	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720		
DB	922	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	981		
QY	721	TAAATTAACAATAAATTTAGAAATTTCCCACTGCAATTTAGGACACTCTCAAACCTTAAA	780		
DB	982	TAAATTAACAATAAATTTAGAAATTTCCCACTGCAATTTAGGACACTCTCAAACCTTAAA	1041		
QY	781	GAACTAGGATTTATAGCAACAATATCAGCTCGATACCTGGAAGCAATTTGTTAGGCAAC	840		
DB	1042	GAACTAGGATTTATAGCAACAATATCAGCTCGATACCTGGAAGCAATTTGTTAGGCAAC	1101		
QY	841	CCTTCTCTTTTACAAATACATTTCTATGACAATCCATCCAAATTTGTTGGAGATCTGCT	900		
DB	1102	CCTTCTCTTTTACAAATACATTTCTATGACAATCCATCCAAATTTGTTGGAGATCTGCT	1161		
QY	901	TTTCAACATTTACCTGAACCTAGAAACCTGCAATCTGAAATGCTGCTCAACAAATCTGAA	960		
DB	1162	TTTCAACATTTACCTGAACCTAGAAACCTGCAATCTGAAATGCTGCTCAACAAATCTGAA	1221		
QY	961	TTTCTGATTTTAACTGGAACCTGCAACCTGGAGCTGCTGCTTAACTGGAGCAGATC	1020		
DB	1222	TTTCTGATTTTAACTGGAACCTGCAACCTGGAGCTGCTGCTTAACTGGAGCAGATC	1281		

721 TTAAATTAACAATTAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
721 TTAAATTAACAATTAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
781 GAACTAGGATTTTCATAGCAACAATATACAGTGCATACCTGAGAAAGCATTTGTAGGCAAC 840
781 GAACTAGGATTTTCATAGCAACAATATACAGTGCATACCTGAGAAAGCATTTGTAGGCAAC 840
841 CCTTCTCTTAATTAACAATATTTCTATGACAAATCCCAATCCCAATTTGTTGGAGATCTGCT 900
841 CCTTCTCTTAATTAACAATATTTCTATGACAAATCCCAATCCCAATTTGTTGGAGATCTGCT 900
901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTTCAAAAATAAATGAA 960
901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAATGGTGGCTTCAAAAATAAATGAA 960
961 TTTCTCTGATTTAACTGGAATCGAAACCTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
961 TTTCTCTGATTTAACTGGAATCGAAACCTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAAATCTCCAAAGTGTAGATCTGTCT 1080
1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAAATCTCCAAAGTGTAGATCTGTCT 1080
1081 TACAACCTATTAAGAAATTTACCCAGTTTTTCAAGTCTGCAAAAGCTTCAAAAATGAC 1140
1081 TACAACCTATTAAGAAATTTACCCAGTTTTTCAAGTCTGCAAAAGCTTCAAAAATGAC 1140
1141 CTAAGACATTAATGAATTAAGTAAATTAAGTGAACATTTCCAGCAGTGTCTAGCCCTC 1200
1141 CTAAGACATTAATGAATTAAGTAAATTAAGTGAACATTTCCAGCAGTGTCTAGCCCTC 1200
1201 CGATCCCTGAAATTTGGCTTGGAAACAAATTTGCTATTATTTACCCCAAGTCAATTTCCACT 1260
1201 CGATCCCTGAAATTTGGCTTGGAAACAAATTTGCTATTATTTACCCCAAGTCAATTTCCACT 1260
1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTGCTTTTCTTAATCT 1320
1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTGCTTTTCTTAATCT 1320
1321 GGGTTACATGGTTTAACTCACTTAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAGG 1380
1321 GGGTTACATGGTTTAACTCACTTAAATTAAGGAAATTAAGGAAATTAAGGAAATTAAGG 1380
1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTCTGT 1440
1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTCTGT 1440
1441 GCAATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAAATCAATGGAATTAAGGTTGACAA 1500
1441 GCAATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAAATCAATGGAATTAAGGTTGACAA 1500
1501 AGCAGTATGAGCAGCTTCAATAGAAAGTGTGGAATGCTTCAAGTCTCAAGTGAACGT 1560
1501 AGCAGTATGAGCAGCTTCAATAGAAAGTGTGGAATGCTTCAAGTCTCAAGTGAACGT 1560
1561 GACCTTGAAGATTTCTGCTGCTTGAAGTGTGAGAAAGTGTGGAATGCTTCAAGTGAACGT 1620
1561 GACCTTGAAGATTTCTGCTGCTTGAAGTGTGAGAAAGTGTGGAATGCTTCAAGTGAACGT 1620
1621 TGTTCACTTCCAGGCCCCCTTCAAAACCTGTAAGACCTGCTTGAATGGCTGGCTGATC 1680
1621 TGTTCACTTCCAGGCCCCCTTCAAAACCTGTAAGACCTGCTTGAATGGCTGGCTGATC 1680
1681 AGAATTTGGAGTGTGAGAAATGCTTGAAGTGTGAGAAATGCTTGAAGTGTGAGAAATGCT 1740
1681 AGAATTTGGAGTGTGAGAAATGCTTGAAGTGTGAGAAATGCTTGAAGTGTGAGAAATGCT 1740
1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCAATTAAGTGTGAAATGCGGCTCATCGCA 1800
1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCAATTAAGTGTGAAATGCGGCTCATCGCA 1800

1801 GCAGTGAAATGCTCTCAAGGAGTCTCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1801 GCAGTGAAATGCTCTCAAGGAGTCTCCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1861 TTTTGCAGCTTTTGCAGCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
1861 TTTTGCAGCTTTTGCAGCAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
1921 GGTGTTTTGTCATTTTGTCTTTCAGAAATCATGTTTCTGCTTCTGCTTCTGCTTCTGCTGCT 1980
1921 GGTGTTTTGTCATTTTGTCTTTCAGAAATCATGTTTCTGCTTCTGCTTCTGCTTCTGCTGCT 1980
1981 GAGCGTGGTCTCTGCTGAAATATTTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
1981 GAGCGTGGTCTCTGCTGAAATATTTCTGCAAAATTTTGAACGAAAGCTCCATTTTCTAGC 2040
2041 CTGAAAGTAAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2041 CTGAAAGTAAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
2101 CTGGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
2101 CTGGGTGGCAGCAAGTATGGCGCTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
2161 AGCACAATGGGCTACATGCTGCTCTCATCTTGTCTCAATTTCCCTTGTCTTCTCATGATG 2220
2161 AGCACAATGGGCTACATGCTGCTCTCATCTTGTCTCAATTTCCCTTGTCTTCTCATGATG 2220
2221 ACCATTGGCTACACCAAGCTCTACCTGCAATTTGGAAAGGAGAGCTGAGAAATATTTGG 2280
2221 ACCATTGGCTACACCAAGCTCTACCTGCAATTTGGAAAGGAGAGCTGAGAAATATTTGG 2280
2281 GACTGCTCTATGGTAAAAACATTTGCCCTTGTGCTCTTCCACCAACTGCACTCTTAACTGCT 2340
2281 GACTGCTCTATGGTAAAAACATTTGCCCTTGTGCTCTTCCACCAACTGCACTCTTAACTGCT 2340
2341 CCGTGGGCTTTCTGCTGCTTCTCTCTTAAATAAACCTTATCAATTTATCACTGCTGAAATG 2400
2341 CCGTGGGCTTTCTGCTGCTTCTCTCTTAAATAAACCTTATCAATTTATCACTGCTGAAATG 2400
2401 ATTAAAGTTTATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
2401 ATTAAAGTTTATCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
2461 ATCTTGTTCATCTCTCACTTTTAAAGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2461 ATCTTGTTCATCTCTCACTTTTAAAGGAGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2520
2521 TGGACAAGATCAAAACACCCAAAGCTTGAATGCTCAATTTAACTCTGATGATGCTGCAAAAACAG 2580
2521 TGGACAAGATCAAAACACCCAAAGCTTGAATGCTCAATTTAACTCTGATGATGCTGCAAAAACAG 2580
2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTG 2640
2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCACTTATGACCTG 2640
2641 CCTCCAGTTCGCTGCCATCACCACTTATCCAGTGACTGAGAGCTGCGCATCTTCTCTCT 2700
2641 CCTCCAGTTCGCTGCCATCACCACTTATCCAGTGACTGAGAGCTGCGCATCTTCTCTCT 2700
2701 GTGGCAATTTGCTCCATGCTC 2721
2701 GTGGCAATTTGCTCCATGCTC 2721

RESULT 15
AAA30779

ID AAA30779 standard; DNA; 2724 BP.

XX AAA30779;

XX 21-AUG-2000 (first entry)

XX

Db 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCAACCTCTCTGCTCTTTTCTCTATACT 1320
Qy 1321 GGGTTACATGGTTAACTCACTTTAAATAAAGGAAATCATGCCCTTACAGAGCTTTGATA 1380
Db 1321 GGGTTACATGGTTAACTCACTTTAAATAAAGGAAATCATGCCCTTACAGAGCTTTGATA 1380
Qy 1381 TCATCTGAAACCTTTCCAGNACTCAAGGTATAGAAATGCGCTTATGCTTACAGAGTCTGT 1440
Db 1381 TCATCTGAAACCTTTCCAGNACTCAAGGTATAGAAATGCGCTTATGCTTACAGAGTCTGT 1440
Qy 1441 GCATTTGGAGTGTGAGAAATGCTATAGAAATTTCTAATCAATGGAATAAAGGTGACAAC 1500
Db 1441 GCATTTGGAGTGTGAGAAATGCTATAGAAATTTCTAATCAATGGAATAAAGGTGACAAC 1500
Qy 1501 AGCAGTATGGACACCTTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGGACACCTTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTGAGTGCAG 1620
Db 1561 GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTGAGTGCAG 1620
Qy 1621 TGTTCACCTTTCCCGAGCCCTTCAAAACCTGTGAAACACCTGCTTCATGSGCTGGTGCATC 1680
Db 1621 TGTTCACCTTTCCCGAGCCCTTCAAAACCTGTGAAACACCTGCTTCATGSGCTGGTGCATC 1680
Qy 1681 AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1740
Db 1681 AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1740
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGGTCACTGCA 1800
Db 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAATTTGGGGTCACTGCA 1800
Qy 1801 GCAGTGAAACATGCTCAGGGAGTCTCAGTGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1801 GCAGTGAAACATGCTCAGGGAGTCTCAGTGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Qy 1861 TTTGGCAGCTTTGACAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 TTTGGCAGCTTTGACAGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy 1921 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
Qy 1981 GAGCGGGTCTCTCTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGGGTCTCTCTGAAATATCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 2041 CTGAAAGTAAATCATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2041 CTGAAAGTAAATCATTTTGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Qy 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Qy 2161 AGCACCATGGCTACATGCTGCTCTCATCTGCTCAATTTCCCTTTGCTTCTCTCATGATG 2220
Db 2161 AGCACCATGGCTACATGCTGCTCTCATCTGCTCAATTTCCCTTTGCTTCTCTCATGATG 2220
Qy 2221 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTTGGCAAGGAGACCTGGAGAAATTTGG 2280
Db 2221 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTTGGCAAGGAGACCTGGAGAAATTTGG 2280
Qy 2281 GACTGCTCTATGTGTAACAACATGCGCTGCTCTTCAACCACTGCAATTTCTAAACTGC 2340
Db 2281 GACTGCTCTATGTGTAACAACATTTGCGCTGCTCTTCAACCACTGCAATTTCTAAACTGC 2340
Qy 2341 CCTGTGCGCTTCTGCTGCTCTCTCTTAAATAAACCCTTACATTTTATCATGCTGCTGAAGTA 2400
Db 2341 CCTGTGCGCTTCTGCTGCTCTCTCTTAAATAAACCCTTATCATGCTGCTGAAGTA 2400

Qy 2401 ATTAAGTTTATCCTTCTGCTGCTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
Db 2401 ATTAAGTTTATCCTTCTGCTGCTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
Qy 2461 ATCTTTGTTCAATCCTCACTTTTAAAGGAGATCTGCTGAGCTGAGAAAGCAAACTTACGTC 2520
Db 2461 ATCTTTGTTCAATCCTCACTTTTAAAGGAGATCTGCTGAGCTGAGAAAGCAAACTTACGTC 2520
Qy 2521 TGGACAAGATCAAAAACACCAAGCTTGTATGCTCAATTAACCTCTGATGATGTCGAAAAACAG 2580
Db 2521 TGGACAAGATCAAAAACACCAAGCTTGTATGCTCAATTAACCTCTGATGATGTCGAAAAACAG 2580
Qy 2581 TCTGTGACTCAACTCAAGCTTGTAAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
Db 2581 TCTGTGACTCAACTCAAGCTTGTAAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
Qy 2641 CCTCCAGTTCCGTCGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCCTCT 2700
Db 2641 CCTCCAGTTCCGTCGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCCTCT 2700
Qy 2701 GTGGCATTTGTCCCATGCTCTTAA 2724
Db 2701 GTGGCATTTGTCCCATGCTCTTAA 2724

Search completed: July 12, 2005, 21:29:18
Job time : 1423 secs

This Page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2724	100.0	2724	4	US-09-170-496D-263	Sequence 263, App
2	2724	100.0	4570	4	US-09-976-594-201	Sequence 201, App
3	2622	96.3	2724	4	US-09-170-496D-277	Sequence 277, App
4	699	25.7	723	4	US-09-495-050A-220	Sequence 220, App
5	25	0.9	25	4	US-09-170-496D-261	Sequence 261, App
6	24	0.9	30	4	US-09-170-496D-260	Sequence 260, App
7	21	0.8	2665	3	US-08-971-089-5	Sequence 5, Appli
8	21	0.8	2896	4	US-09-266-225D-9	Sequence 9, Appli
9	21	0.8	2906	4	US-09-949-016-669	Sequence 669, App
10	21	0.8	2910	4	US-09-949-016-4412	Sequence 4412, Ap
11	21	0.8	22339	4	US-09-949-016-12411	Sequence 12411, A
12	21	0.8	22339	4	US-09-949-016-16154	Sequence 16154, A
13	21	0.8	61198	4	US-09-949-016-17248	Sequence 17248, A
14	20	0.7	443	4	US-09-821-976-13818	Sequence 13818, A
15	20	0.7	574	4	US-09-270-767-14190	Sequence 14190, A
16	20	0.7	6101	4	US-09-949-016-139111	Sequence 139111,
17	20	0.7	2148	4	US-09-949-016-2145	Sequence 2145, Ap
18	20	0.7	2659	4	US-09-902-540-4695	Sequence 4695, Ap
19	20	0.7	28558	4	US-09-902-540-1231	Sequence 1231, Ap
20	20	0.7	68529	4	US-09-949-016-12140	Sequence 12140, A
21	20	0.7	68529	4	US-09-949-016-15671	Sequence 15671, A
22	19	0.7	601	4	US-09-949-016-29857	Sequence 29857, A
23	19	0.7	601	4	US-09-949-016-55155	Sequence 55155, A
24	19	0.7	601	4	US-09-949-016-134703	Sequence 134703,
25	19	0.7	601	4	US-09-949-016-134704	Sequence 134704,
26	19	0.7	601	4	US-09-949-016-134705	Sequence 134705,
27	19	0.7	601	4	US-09-949-016-153352	Sequence 153352,

Db 361 |||||CTGCAGAAATACAGCTAAGACACAGTACCCACAGAAAGCTCTGCAGAAATTTGGAGACCTT 420
Qy 421 CAATCCCTCGGCTCGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 421 CAATCCCTCGGCTCGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Qy 481 CTGCATTCCTCGAGGACCTGTGGCTGGATGACAAATGGTTTAAACAGAAATCCCCTGCAG 540
Db 481 CTGCATTCCTCGAGGACCTGTGGCTGGATGACAAATGGCTTTAAACAGAAATCCCCTGCAG 540
Qy 541 GCTTTTAGAAGTTTATCGGCATTCGAAGCCATGACCTTTGGCCCTGAAACAAAATACACCAC 600
Db 541 GCTTTTAGAAGTTTATCGGCATTCGACGATGACCTTTGGCCCTGAAACAAAATACACCAC 600
Qy 601 ATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 601 ATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 721 TTAATAATTACAATTAACCTTTCCCACTGCAATTTAGGACACTCTCCAAACCTTAAA 780
Db 721 TTAATAATTACAATTAACCTTTCCCACTGCAATTTAGGACACTCTCCAAACCTTAAA 780
Qy 781 GAACTAGGATTTTATPAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Qy 841 CCTTCTCTTATACAATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 841 CCTTCTCTTATACAATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Qy 901 TTTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCCTCACAATAACTGAA 960
Db 901 TTTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCCTCACAATAACTGAA 960
Qy 961 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 961 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCCAGTCTAGATCTGTCT 1080
Db 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCCAGTCTAGATCTGTCT 1080
Qy 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Qy 1141 CTAAGACATAATGAAATCTAGAAAATTAAGTTGACATTTTCAGCAGTGTGCTTAGACCTC 1200
Db 1141 CTAAGACATAATGAAATCTAGAAAATTAAGTTGACATTTTCAGCAGTGTGCTTAGACCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTACCCCAATGCAATTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTACCCCAATGCAATTTTCCTACT 1260
Qy 1261 TTGGCAATCCCTAATAAAGCTGACCTATCGTCCAACTCTGTCGTCTTTTCCCTATAACT 1320
Db 1261 TTGGCAATCCCTAATAAAGCTGACCTATCGTCCAACTCTGTCGTCTTTTCCCTATAACT 1320
Qy 1321 GGGTTACATGGTTTAACTCACTTAAATTTAAACAGAAATTCATGCTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGGTTTAACTCACTTAAATTTAAACAGAAATTCATGCTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAAACCTTTCCAGAACTCAAGTTTATAGAAATGCCCTTATGCTTACAGTGTGT 1440
Db 1381 TCATCTGAAACCTTTCCAGAACTCAAGTTTATAGAAATGCCCTTATGCTTACAGTGTGT 1440
Qy 1441 GCATTTGGAGTGTGTGAGAAATCCCTTATAAGATTTCTAATCAATGGAATAAAGGTGACCAAC 1500

Db 1441 GCATTTGGAGTGTGTGAGAAATGCCCTATAAGATTTCTAATCAATGGAATAAAGGTGACCAAC 1500
Qy 1501 AGCAGTATGACGACCTTTATAGAAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGACGACCTTTATAGAAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTGAAGATTTCTGCTTGGCTTTGAGGAACACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db 1561 GACCTTGAAGATTTCTGCTTGGCTTTGAGGAACACCTGAAAGCCCTTCATTCAGTGCAG 1620
Qy 1621 TGTTTCACTTTCCAGGCCCTTCAAAACCTGTGAAACACCTGTGTGATGCTGCTGATC 1680
Db 1621 TGTTTCACTTTCCAGGCCCTTCAAAACCTGTGAAACACCTGTGTGATGCTGCTGATC 1680
Qy 1681 AGAATTGGAGTGTGGACCATAGAGATTTCTGGCAGTTCTGACCTTCTGATGGCTGATC 1740
Db 1681 AGAATTGGAGTGTGGACCATAGAGATTTCTGGCAGTTCTGACCTTCTGATGGCTGATC 1740
Qy 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATAAACTGTTAAATTTGGGGTCAATCGCA 1800
Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCAATAAACTGTTAAATTTGGGGTCAATCGCA 1800
Qy 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTSCCGTGTGCTGTGGATGCGTTCACT 1860
Db 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTSCCGTGTGCTGTGGATGCGTTCACT 1860
Qy 1861 TTTGGGAGCTTTGACGACATGCTGCTGTGGGAGAAATGGGGTGGTGGCAATGTCATTT 1920
Db 1861 TTTGGGAGCTTTGACGACATGCTGCTGTGGGAGAAATGGGGTGGTGGCAATGTCATTT 1920
Qy 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTTGGCAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 2041 CTGAAAGTAAATCATTTTGTCTGTGCCCTGTGCTGCTTGACCATGGCGGAGTTCCCTCG 2100
Db 2041 CTGAAAGTAAATCATTTTGTCTGTGCCCTGTGCTGCTTGACCATGGCGGAGTTCCCTCG 2100
Qy 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGGCC 2160
Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGGCC 2160
Qy 2161 AGCACCATGGGCTACATGCTGCTCATCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Db 2161 AGCACCATGGGCTACATGCTGCTCATCTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Qy 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATTTGG 2280
Db 2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATTTGG 2280
Qy 2281 GACTGCTCTATGTTAAACACATTTGCCCTGTGCTCTTCAACCACTGCATCTCTAACTGC 2340
Db 2281 GACTGCTCTATGTTAAACACATTTGCCCTGTGCTCTTCAACCACTGCATCTCTAACTGC 2340
Qy 2341 CCTGTGGCTTTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAATG 2400
Db 2341 CCTGTGGCTTTTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAATG 2400
Qy 2401 ATTAAGTTTATCTTCTGTTGGTAGTCCCACTCTGCAATGCTCAATCCCCTCTCTAC 2460
Db 2401 ATTAAGTTTATCTTCTGTTGGTAGTCCCACTCTGCAATGCTCAATCCCCTCTCTAC 2460
Qy 2461 ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGAGCCTTGAGAAAACCACTACGTC 2520
Db 2461 ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGAGCCTTGAGAAAACCACTACGTC 2520
Qy 2521 TGGACAAGATCAAAACACCCCAAGCTTGAATGCAATTAATCTCTGATGATGTCGAAAAACAG 2580
Db 2521 TGGACAAGATCAAAACACCCCAAGCTTGAATGCAATTAATCTCTGATGATGTCGAAAAACAG 2580

[illegible]

RESULT 4
US-09-495-050A-220
; Sequence 220, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 220
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2603450CT1
US-09-495-050A-220

Query Match 25.7%; Score 699; DB 4; Length 723;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 699; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1695 TTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGAATGCTTTGGTGTGACTTCAACAG 1744
Db 1 TTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGAATGCTTTGGTGTGACTTCAACAG 60

QY 1745 TTTTCAGATCCCTCTGTACATTTCCCCCATTAATACTGTTAATGGGGTCAATCGCAGCAG 1804
Db 61 TTTTCAGATCCCTCTGTACATTTCCCCCATTAATACTGTTAATGGGGTCAATCGCAGCAG 120

QY 1805 TGAACATGCTCAGGAGTCTCCAGTGGCGTCTGGCTGTGTGTGATGCGTTCACTTTTG 1864
Db 121 TGAACATGCTCAGGAGTCTCCAGTGGCGTCTGGCTGTGTGTGATGCGTTCACTTTTG 180

QY 1865 GCAGCTTTGCAGCATGTGCTGCTGGTGGAGAAATGGGTTGGTGGCCATGTCATTGGTT 1924
Db 181 GCAGCTTTGCAGCATGTGCTGCTGGTGGAGAAATGGGTTGGTGGCCATGTCATTGGTT 240

QY 1925 TTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGTTTCTGTTACTCTGGCAGCCCTGGAGC 1984
Db 241 TTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGTTTCTGTTACTCTGGCAGCCCTGGAGC 300

QY 1985 GTGGGTTCTCTGAAATATCTGAAATATTTGAAACGAAAGCTCCATTTCTAGCCTGA 2044
Db 301 GTGGGTTCTCTGAAATATCTGAAATATTTGAAACGAAAGCTCCATTTCTAGCCTGA 360

QY 2045 AAGTAATCATTTTGTCTGTGCGCTGCTGGCCCTTGACCATGCGCAGTTCCTCCCTGCTGG 2104
Db 361 AAGTAATCATTTTGTCTGTGCGCTGCTGGCCCTTGACCATGCGCAGTTCCTCCCTGCTGG 420

QY 2105 GTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTGCGGAGCCCGCAGCA 2164
Db 421 GTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTGCGGAGCCCGCAGCA 480

QY 2165 CCATGGGCTACATGTCGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATGACCA 2224
Db 481 CCATGGGCTACATGTCGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATGACCA 540

QY 2225 TTGCGCTACACCAAGCTCTACTGCAATTTGGAACGGGAGCCTGGAGATATTTGGGACT 2284
Db 541 TTGCGCTACACCAAGCTCTACTGCAATTTGGAACGGGAGCCTGGAGATATTTGGGACT 600

QY 2285 GCTCTATGTAACCAACATGTCCTGCTCTTCAACCACTGCATCCTTAAACCTGCCCTG 2344
Db 2285 GCTCTATGTAACCAACATGTCCTGCTCTTCAACCACTGCATCCTTAAACCTGCCCTG 2344

Db 601 GCTCTATGTAACCAACATGTCCTGCTCTTCAACCACTGCATCCTTAAACCTGCCCTG 660

QY 2345 TGGCTTTTCTGTCCTCTCTCTCTTTTAATAAACCTTACAT 2383
Db 661 TGGCTTTTCTGTCCTCTCTCTCTTTTAATAAACCTTACAT 699

RESULT 5
US-09-170-496D-261
; Sequence 261, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 261
; LENGTH: 25
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. 6555339el Sequence
US-09-170-496D-261

Query Match 0.9%; Score 25; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.021;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1313 CTATAACTGGGTTACATGGTTTAAC 1337
Db 1 CTATAACTGGGTTACATGGTTTAAC 25

RESULT 6
US-09-170-496D-260/c
; Sequence 260, Application US/09170496D
; Patent No. 6555339
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. 6555339-Endogenous, Constitutively Activated Human G Protein-
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/09/170,496D
; CURRENT FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 260
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. 6555339el Sequence
US-09-170-496D-260

Query Match 0.9%; Score 24; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.07;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1426 GCTTACCAGTGTGTGCAATTGGA 1449
Db 30 GCTTACCAGTGTGTGCAATTGGA 7

RESULT 7
US-08-971-089-5


```
; Sequence 5, Application US/08971089
; Patent No. 6376174
; GENERAL INFORMATION:
; APPLICANT: Pulist, Stefan M.
; APPLICANT: Scoles, Daniel R.
; TITLE OF INVENTION: NUCLEIC ACID ENCODING
; TITLE OF INVENTION: SCHWANNIN-BINDING-PROTEINS AND PRODUCTS RELATED THERETO
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CAMPBELL & FLORES, LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,089
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,987
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Ramos, Robert T.
; REGISTRATION NUMBER: 37,915
; REFERENCE/DOCKET NUMBER: P-CE 2862
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)535-9001
; TELEFAX: (619)535-8949
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: both
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 5..2077
; US-08-971-089-5

Query Match 0.8%; Score 21; DB 3; Length 2665;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATGGGCTACATGG 2179
Db 1815 CCAGCACCATGGGCTACATGG 1835

RESULT 8
US-09-266-225D-9
; Sequence 9, Application US/09266225D
; Patent No. 6573364
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishan
; APPLICANT: Kingsmore, Stephen
; APPLICANT: Tchernov, Velizar
; TITLE OF INVENTION: Isolation and Characterization of Hermansky-Pudlak
; TITLE OF INVENTION: Syndrome (HPS) Protein Complexes and HPS Protein-
; TITLE OF INVENTION: Interacting Proteins
; FILE REFERENCE: 15966-523
; CURRENT APPLICATION NUMBER: US/09/266,225D
; CURRENT FILING DATE: 1999-03-10
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2896
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-266-225D-9

Query Match 0.8%; Score 21; DB 4; Length 2896;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATGGGCTACATGG 2179
Db 2132 CCAGCACCATGGGCTACATGG 2152

RESULT 9
US-09-949-016-669
; Sequence 669, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 669
; LENGTH: 2906
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-669

Query Match 0.8%; Score 21; DB 4; Length 2906;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCATGGGCTACATGG 2179
Db 2147 CCAGCACCATGGGCTACATGG 2167

RESULT 10
US-09-949-016-4412
; Sequence 4412, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4412
; LENGTH: 2910
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-4412

Query Match 0.8%; Score 21; DB 4; Length 2910;
```



```
Best Local Similarity 100.0%; Pred. No. 3.6; DB 4; Length 22339;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCACATGGGCTACATGG 2179
      |||||
Db 2151 CCAGCACCACATGGGCTACATGG 2171

RESULT 11
US-09-949-016-12411
; Sequence 12411, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12411
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(22339)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12411

Query Match 0.8%; Score 21; DB 4; Length 22339;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCACATGGGCTACATGG 2179
      |||||
Db 18974 CCAGCACCACATGGGCTACATGG 18994

RESULT 12
US-09-949-016-16154
; Sequence 16154, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16154
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(22339)
; OTHER INFORMATION: n = A,T,C or G
```

```
US-09-949-016-16154

Query Match 0.8%; Score 21; DB 4; Length 22339;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2159 CCAGCACCACATGGGCTACATGG 2179
      |||||
Db 18974 CCAGCACCACATGGGCTACATGG 18994

RESULT 13
US-09-949-016-17248
; Sequence 17248, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17248
; LENGTH: 61198
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(61198)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17248

Query Match 0.8%; Score 21; DB 4; Length 61198;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1994 CTGTGAAATATTTCTGCAAAAT 2014
      |||||
Db 35060 CTGTGAAATATTTCTGCAAAAT 35080

RESULT 14
US-09-621-976-13818/c
; Sequence 13818, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13818
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 405
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-13818
```


Query Match 0.7%; Score 20; DB 4; Length 443;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 255 CCTGCCCGAGTCCTCCGTTCC 274
|||||
Db 315 CCTGCCCGAGTCCTCCGTTCC 296

RESULT 15

US-09-270-767-14190
; Sequence 14190, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14190
; LENGTH: 574
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14190

Query Match 0.7%; Score 20; DB 4; Length 574;
Best Local Similarity 100.0%; Pred.No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 715 TTAGATTAAATTACAATAA 734
|||||
Db 415 TTAGATTAAATTACAATAA 434

Search completed: July 13, 2005, 03:13:17
Job time : 457 secs

This Page Blank (uspto)

1	2724	100.0	2724	15	US-10-351-385-263	Sequence 263, App
2	2724	100.0	2724	15	US-10-325-567A-431	Sequence 421, App
3	2724	100.0	2724	17	US-10-174-456-A	Sequence 4, Appl
4	2724	100.0	2724	20	US-10-751-736-B	Sequence 21, Appl
5	2724	100.0	2880	17	US-10-395-027-1114	Sequence 1114, App
6	2724	100.0	2880	21	US-10-482-029-157	Sequence 157, App
7	2724	100.0	2880	21	US-10-651-237-48	Sequence 48, Appl

103

Constitutively Activated Human G

|||||
241 Db CTGCTCCGAATCCCCTGCCAGTCTCCGCTTCTTGAGGAGTTAGCTTTGGGGAAC 300
|||||
301 QY GCTCTGACATACATTCCTCAAGGAGCATTCACCTGGCTTTACAGTCTTAAAGTTCCTATG 360
|||||
301 Db GCTCTGACATACATTCCTCAAGGAGCATTCACCTGGCTTTACAGTCTTAAAGTTCCTATG 360
|||||
361 QY CTGCGAATAATCAGCTAAGACACACGTAACCAAGAGCTCTGAGAAATTTGGGAAGCCTT 420
361 Db CTGCGAATAATCAGCTAAGACACACGTAACCAAGAGCTCTGAGAAATTTGGGAAGCCTT 420
|||||
421 QY CAATCCCTGGCTGCTGATGCTAACACACATGATGTCGCCCAAGCTGTTTCAGTGGC 480
421 Db CAATCCCTGGCTGCTGATGCTAACACACATGATGTCGCCCAAGCTGTTTCAGTGGC 480
|||||
481 QY CTGCATTCCTCTGAGGCACCTGTGGCTGGATGACAATGCGTTAAACAGAAATCCCGTCCAG 540
481 Db CTGCATTCCTCTGAGGCACCTGTGGCTGGATGACAATGCGTTAAACAGAAATCCCGTCCAG 540
|||||
541 QY GCTTTTGAAGTTTATCGGCATTTGCAAGCATGACCTTGCCCTGAAACAAATACACCAAC 600
541 Db GCTTTTGAAGTTTATCGGCATTTGCAAGCATGACCTTGCCCTGAAACAAATACACCAAC 600
|||||
601 QY ATACACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAT 660
601 Db ATACACAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAT 660
|||||
661 QY AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGCTCCACAGCTTAGAGACTTTAGAT 720
661 Db AGAATCCACTCCCTGGGAAGAAATGCTTTGATGGCTCCACAGCTTAGAGACTTTAGAT 720
|||||
721 QY TTAATAATCAATAACCTTGATGAATTTCCCACTGCAATTTAGCACTCTCCAACTTAAA 780
721 Db TTAATAATCAATAACCTTGATGAATTTCCCACTGCAATTTAGCACTCTCCAACTTAAA 780
|||||
781 QY GAACCTAGGATTTTATAGCAACAATATCAGTGTGATACCTTGAGAAAGCATTTGTAGGCAAC 840
781 Db GAACCTAGGATTTTATAGCAACAATATCAGTGTGATACCTTGAGAAAGCATTTGTAGGCAAC 840
|||||
841 QY CTTCTCTTATTACATAATATTTCTATGACAATCCCAATTTGTTGGGAGATCTGCT 900
841 Db CTTCTCTTATTACATAATATTTCTATGACAATCCCAATTTGTTGGGAGATCTGCT 900
|||||
901 QY TTTCAACATTTTAACTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAAATGAA 960
901 Db TTTCAACATTTTAACTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAAATGAA 960
|||||
961 QY TTTCTGATTTTAACTGGAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
961 Db TTTCTGATTTTAACTGGAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
|||||
1021 QY TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGCT 1080
1021 Db TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGCT 1080
|||||
1081 QY TACAACCTATTAGAAGATTTACCCAGTCTGCACTGCGCAAAAGCTTCAGAAAATTGAC 1140
1081 Db TACAACCTATTAGAAGATTTACCCAGTCTGCACTGCGCAAAAGCTTCAGAAAATTGAC 1140
|||||
1141 QY CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACATTTCCAGCAGTTGCTTAGGCCTC 1200
1141 Db CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACATTTCCAGCAGTTGCTTAGGCCTC 1200
|||||
1201 QY CGATCCCTGAAATTTGGCTTGGACAAAATTTGCTATTATTATTCACCCCAATGCTTTCCACT 1260
1201 Db CGATCCCTGAAATTTGGCTTGGACAAAATTTGCTATTATTATTCACCCCAATGCTTTCCACT 1260
|||||
1261 QY TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCTCTATAACT 1320
1261 Db TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCTCTATAACT 1320
|||||
1321 QY GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
|||||

1321 Db GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
|||||
1381 QY TCATCTGAAAACTTTCCAGAACTCAAGGTATAGAAATGCTTATGCTTTACAGTGTGT 1440
|||||
1381 Db TCATCTGAAAACTTTCCAGAACTCAAGGTATAGAAATGCTTATGCTTTACAGTGTGT 1440
|||||
1441 QY GCATTTGAGTGTGTGAGAAATGCCCTATAAGATTTCCTAATCAATGGAATAAAGGTGACAA 1500
1441 Db GCATTTGAGTGTGTGAGAAATGCCCTATAAGATTTCCTAATCAATGGAATAAAGGTGACAA 1500
|||||
1501 QY AGCAGTATGAGACACTTTTCAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
1501 Db AGCAGTATGAGACACTTTTCAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
|||||
1561 QY GACCTTTGAAGATTTCTCTGCTTTGAGGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
1561 Db GACCTTTGAAGATTTCTCTGCTTTGAGGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
|||||
1621 QY TGTTCACCTTCCCAGAGGCCCTTCAAACTGTGTGAACACCTGCTTGTATGGCTGGCTGATC 1680
1621 Db TGTTCACCTTCCCAGAGGCCCTTCAAACTGTGTGAACACCTGCTTGTATGGCTGGCTGATC 1680
|||||
1681 QY AGAATTTGAGTGTGTGAGCAATAGCAGTTCTGGCACCTTCTGTAATGCTTTGGTGACTTCA 1740
1681 Db AGAATTTGAGTGTGTGAGCAATAGCAGTTCTGGCACCTTCTGTAATGCTTTGGTGACTTCA 1740
|||||
1741 QY ACAGTTTTTCCAGATCCCTCTGTACATTTTCCCCCATTTAAACTGTTAAATTTGGGTCATCGCA 1800
1741 Db ACAGTTTTTCCAGATCCCTCTGTACATTTTCCCCCATTTAAACTGTTAAATTTGGGTCATCGCA 1800
|||||
1801 QY GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCTGCTGGTGTGTGATGCGGTTCATC 1860
1801 Db GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCTGCTGGTGTGTGATGCGGTTCATC 1860
|||||
1861 QY TTTGGCAGCTTTGACGACATGCTGCTGGTGGGAGAAATGGGTTGGTTCATGCTCAT 1920
1861 Db TTTGGCAGCTTTGACGACATGCTGCTGGTGGGAGAAATGGGTTGGTTCATGCTCAT 1920
|||||
1921 QY GGTTTTTTGTCCATTTTGTCTTCAAAATCATCTGTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
1921 Db GGTTTTTTGTCCATTTTGTCTTCAAAATCATCTGTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
|||||
1981 QY GAGCGTGGTTCCTGTGAAAATATTTCTGCAAAAATTTGAAAACGAAAGCTCCATTTCTAGC 2040
1981 Db GAGCGTGGTTCCTGTGAAAATATTTCTGCAAAAATTTGAAAACGAAAGCTCCATTTCTAGC 2040
|||||
2041 QY CTGAAAGTAAATCATTTTGTCTCTGTCCTGCTGCTTGAACCATGCGCGAGTCCCTCTG 2100
2041 Db CTGAAAGTAAATCATTTTGTCTCTGTCCTGCTGCTTGAACCATGCGCGAGTCCCTCTG 2100
|||||
2101 QY CTGGTGGCAGCAAGTATGGGCGCTCCCTCTCTGCTGCTGCTTTGCTTTTGGGAGGCC 2160
2101 Db CTGGTGGCAGCAAGTATGGGCGCTCCCTCTCTGCTGCTGCTTTGCTTTTGGGAGGCC 2160
|||||
2161 QY AGCACAATGGGCTACATGGTGTCTCATTTGCTCAATTCCTCTTGTCTTCTCATGATG 2220
2161 Db AGCACAATGGGCTACATGGTGTCTCATTTGCTCAATTCCTCTTGTCTTCTCATGATG 2220
|||||
2221 QY ACCATTGCTTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATTTGG 2280
2221 Db ACCATTGCTTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATTTGG 2280
|||||
2281 QY GACTGCTCTATGGTAAAAACATTTGCCCTGTTGCTTTTCCACCACTGCATCTCTAACTGC 2340
2281 Db GACTGCTCTATGGTAAAAACATTTGCCCTGTTGCTTTTCCACCACTGCATCTCTAACTGC 2340
|||||
2341 QY CCGTGGCTTTCTGTCTCTTCTTCTTAAACCTTACATTTATCAGTCTCTGAATG 2400
2341 Db CCGTGGCTTTCTGTCTCTTCTTCTTAAACCTTACATTTATCAGTCTCTGAATG 2400
|||||
2401 QY ATTAAGTTTATCTTCTGCTGGTAGTCCACTTCTGCTGATGCTCAATCCCTCTCTAC 2460
2401 Db ATTAAGTTTATCTTCTGCTGGTAGTCCACTTCTGCTGATGCTCAATCCCTCTCTAC 2460
|||||

QY 61 GGGGCGAGCTCTCCAGGCTGTGTGTGCTGAGGGGCTGCCCAACACACTGTCATTGC 120
DB 109 GGGGCGAGCTCTCCAGGCTGTGTGTGCTGAGGGGCTGCCCAACACACTGTCATTGC 168
QY 121 GAGCCCGACGCGAGGATTTGCTCAGGGTGGACTGCTCCGACCTCTGGGCTCTCGGAGCTG 180
DB 169 GAGCCCGACGCGAGGATTTGCTCAGGGTGGACTGCTCCGACCTCTGGGCTCTCGGAGCTG 228
QY 181 CTTCTCAACCTCAGGCTCTTCACTCTCTACCTAGACCTCAGTATGAACAACATCAGTCAG 240
DB 229 CTTCTCAACCTCAGGCTCTTCACTCTCTACCTAGACCTCAGTATGAACAACATCAGTCAG 288
QY 241 CTGCTCCCGAATCCCTGCGGCTCTCCGCTCTCTGAGGAGTTACGCTTTGCGGGAAC 300
DB 289 CTGCTCCCGAATCCCTGCGGCTCTCCGCTCTCTGAGGAGTTACGCTTTGCGGGAAC 348
QY 301 GCTCTGACATACATTTCCCAAGGGAGCAATTCACCTGGCCCTTTACAGTCTTAAAGTTCTTTATG 360
DB 349 GCTCTGACATACATTTCCCAAGGGAGCAATTCACCTGGCCCTTTACAGTCTTAAAGTTCTTTATG 408
QY 361 CTGCAGAAATATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
DB 409 CTGCAGAAATATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 468
QY 421 CAATCCCTGCGCTGCGATGCTAAACACATCAGCTATGCCCCCAAGCTGTTTCAGTGGC 480
DB 469 CAATCCCTGCGCTGCGATGCTAAACACATCAGCTATGCCCCCAAGCTGTTTCAGTGGC 528
QY 481 CTGCAATCCCTGAGGACCTCTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGTCCAG 540
DB 529 CTGCAATCCCTGAGGACCTCTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGTCCAG 588
QY 541 GCTTTTAGAAGTTATTCGGCAATTCGAAGCAATGACCTTGGCCCTGGAACAAATACACCAC 600
DB 589 GCTTTTAGAAGTTATTCGGCAATTCGAAGCAATGACCTTGGCCCTGGAACAAATACACCAC 648
QY 601 ATACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
DB 649 ATACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTCCATAACAAT 708
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
DB 709 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 768
QY 721 TTAATATCAATAACCTTGATGAATTTCCCACTGCAATTAGACACTCTCCAACTTAAA 780
DB 769 TTAATATCAATAACCTTGATGAATTTCCCACTGCAATTAGACACTCTCCAACTTAAA 828
QY 781 GAACTAGGATTTATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB 829 GAACTAGGATTTATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC 888
QY 841 CTTCTCTTATTAACAATACATTTCTATGACAATCCCATCCCAATTTGTTGGAGATCTGCT 900
DB 889 CTTCTCTTATTAACAATACATTTCTATGACAATCCCATCCCAATTTGTTGGAGATCTGCT 948
QY 901 TTTCAACATTTACCTGAACTAAGAACACTGCTGTAATGCTGCTCAACAATTAACCTGAA 1008
DB 949 TTTCAACATTTACCTGAACTAAGAACACTGCTGTAATGCTGCTCAACAATTAACCTGAA 1040
QY 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAAGATC 1020
DB 1009 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAAGATC 1068
QY 1021 TCATCTCTTCTCAAAACCTCTGCAATCAGTTACCTTAATCTCCAAGTCTAGATCTGTCT 1080
DB 1069 TCATCTCTTCTCAAAACCTCTGCAATCAGTTACCTTAATCTCCAAGTCTAGATCTGTCT 1128
QY 1081 TACAACCTTATAGAGATTTACCCAGTTTTCAGTCTGCGCAAGAGCTTCAGAAATTTGAC 1140
DB 1129 TACAACCTTATAGAGATTTACCCAGTTTTCAGTCTGCGCAAGAGCTTCAGAAATTTGAC 1188
QY 1141 CTAAGACATATGAATCTACGAAATTAAGTTGACACTTTTCCAGCAGTTGTTAGCCTC 1200

DB 1189 CTAAGACATATGAATCTAGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1248
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTACACCCCAATGCAATTTCCACT 1260
DB 1249 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTACACCCCAATGCAATTTCCACT 1308
QY 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCCTCGTCTGCTTTTCTATAACT 1320
DB 1309 TTGCCATCCCTAATAAGCTGGACCTATCGTCCAACTCCTCGTCTGCTTTTCTATAACT 1368
QY 1321 GGGTTACATGGTTTAACTCACTTAAATAATTAACAGGAAATCATGCCTTACAGAGCTTGATA 1380
DB 1369 GGGTTACATGGTTTAACTCACTTAAATAATTAACAGGAAATCATGCCTTACAGAGCTTGATA 1428
QY 1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACCAAGTGTCT 1440
DB 1429 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTTATGCTTACCAAGTGTCT 1488
QY 1441 GCATTTGGAGTGTGAGAAATGCTTAACTCAATCAATGGAATAAAGGTGACAAAC 1500
DB 1489 GCATTTGGAGTGTGAGAAATGCTTAACTCAATCAATGGAATAAAGGTGACAAAC 1548
QY 1501 AGCAGTATGACGACCTTTCAATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
DB 1549 AGCAGTATGACGACCTTTCAATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1608
QY 1561 GACCTTGAAGATTTCTGCTTGAATTGAGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
DB 1609 GACCTTGAAGATTTCTGCTTGAATTGAGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1668
QY 1621 TGTTTACCTTTCCCGAGCCCTTCAAACTCTGTAACACCTGCTGTAATGCTGCTGATC 1680
DB 1669 TGTTTACCTTTCCCGAGCCCTTCAAACTCTGTAACACCTGCTGTAATGCTGCTGATC 1728
QY 1681 AGAATTTGAGTGTGGACCATAGCAGTTCTGGCACTTACTTTGTAATGCTTTGGTGAATCA 1740
DB 1729 AGAATTTGAGTGTGGACCATAGCAGTTCTGGCACTTACTTTGTAATGCTTTGGTGAATCA 1788
QY 1741 ACAGTTTTCAGATCCCTCTGTATACATTTCCCACTTAACTGTTAAATTTGGGCTCATCGCA 1800
DB 1789 ACAGTTTTCAGATCCCTCTGTATACATTTCCCACTTAACTGTTAAATTTGGGCTCATCGCA 1848
QY 1801 GCAGTGAACATGCTCAGGAGTCTCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
DB 1849 GCAGTGAACATGCTCAGGAGTCTCCAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908
QY 1861 TTTGGCAGCTTTGCAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
DB 1909 TTTGGCAGCTTTGCAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1968
QY 1921 GGTTTTTTGTCCATTTTGTCTCAGATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 1980
DB 1969 GGTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 2028
QY 1981 GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB 2029 GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2088
QY 2041 CTGAAAGTAAATCAATTTTGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
DB 2089 CTGAAAGTAAATCAATTTTGTCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2148
QY 2101 CTGCGTGGCAGCAGTATGGGCTCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
DB 2149 CTGCGTGGCAGCAGTATGGGCTCTCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2208
QY 2161 AGCACCATGGCTACATGCTGCTCTCATCTTGTCTCAATTTCCCTTTGCTTCTCTCATGATG 2220
DB 2209 AGCACCATGGCTACATGCTGCTCTCATCTTGTCTCAATTTCCCTTTGCTTCTCTCATGATG 2268
QY 2221 ACCATTCCTTACCAACAGCTCTACTGCAATTTTGAAGAAGGAGACCTTGGAGAAATATTGG 2280

Qy	421	CAATCCCTGGGTCTGGAGTCTAACCA	CATCAGCTATGTGCCCCAAGCTGTTTCA	GGC 480
Db	469	CAATCCCTGGGTCTGGAGTCTAACCA	CATCAGCTATGTGCCCCAAGCTGTTTCA	GGC 528
Qy	481	CTGCATTTCCCTGAGGCACCTGTGCT	GGATGACAAATGGTTTACAGAAATCC	CCGTC 540
Db	529	CTGCATTTCCCTGAGGCACCTGTGCT	GGATGACAAATGGTTTACAGAAATCC	CCGTC 588
Qy	541	GCTTTTAGAAAGTTTATCGGCATTC	GCAAGCCATGACCTTTGGCCCTGAA	CAAAATACAC 600
Db	589	GCTTTTAGAAAGTTTATCGGCATTC	GCAAGCCATGACCTTTGGCCCTGAA	CAAAATACAC 648
Qy	601	ATACCAGCATATGCTTTTGGAAACCT	CTCCAGCTTGGTAGTTCTACATCTCC	ATAACAAT 660
Db	649	ATACCAGCATATGCTTTTGGAAACCT	CTCCAGCTTGGTAGTTCTACATCTCC	ATAACAAT 708
Qy	661	AGAAATCCACTCCCTGGGAAAGAAAT	GTGTTGATGGGCTCCACAGCTTAGA	GTATTAGAT 720
Db	709	AGAAATCCACTCCCTGGGAAAGAAAT	GTGTTGATGGGCTCCACAGCTTAGA	GTATTAGAT 768
Qy	721	TTAAATTAACAATAACCTTTGATGAAT	TCCCAATTCAGCACTTAGCACTCTCC	AAACCTTAAA 780
Db	769	TTAAATTAACAATAACCTTTGATGAAT	TCCCAATTCAGCACTTAGCACTCTCC	AAACCTTAAA 828
Qy	781	GAACTAGGATTTTCATAGCAACAATAT	CAGGTTCGATCTGAGAAAGCAATTTG	TAGGCAAC 840
Db	829	GAACTAGGATTTTCATAGCAACAATAT	CAGGTTCGATCTGAGAAAGCAATTTG	TAGGCAAC 888
Qy	841	CCCTCTCTTATTAACAATACATTTCT	ATGACAAATCCCATTCAAATTTGTGG	GAGATCTGCT 900
Db	889	CCCTCTCTTATTAACAATACATTTCT	ATGACAAATCCCATTCAAATTTGTGG	GAGATCTGCT 948
Qy	901	TTTTCAACATTTACCTGAACTAAGAA	CACTGCATCTGAAATGTCCTCACAA	ATAAATGAA 960
Db	949	TTTTCAACATTTACCTGAACTAAGAA	CACTGCATCTGAAATGTCCTCACAA	ATAAATGAA 1008
Qy	961	TTTTCTCTGATTTAACTGGAATTCG	CAAACTCGAGAGTCTGACTTTTAA	CTGGAGCACAGATC 1020
Db	1009	TTTTCTCTGATTTAACTGGAATTCG	CAAACTCGAGAGTCTGACTTTTAA	CTGGAGCACAGATC 1068
Qy	1021	TCATCTCTTCTCAAAACCGTCTGCA	ATTCAGTTACCTAAATCTCAAGTGCT	TAGATCTGTCT 1080
Db	1069	TCATCTCTTCTCAAAACCGTCTGCA	ATTCAGTTACCTAAATCTCAAGTGCT	TAGATCTGTCT 1128
Qy	1081	TACAACCTATTTAGAAAGATTTAC	CCAGTTTTTTCAGTCTGCCAAAGCTT	CAGAAATGAC 1140
Db	1129	TACAACCTATTTAGAAAGATTTAC	CCAGTTTTTTCAGTCTGCCAAAGCTT	CAGAAATGAC 1188
Qy	1141	CTAAGACATAATGAAATCTACGAAAT	TAAAGTTTGACATTTTCCAGCAGTT	TGCTTAGCCCTC 1200
Db	1189	CTAAGACATAATGAAATCTACGAAAT	TAAAGTTTGACATTTTCCAGCAGTT	TGCTTAGCCCTC 1248
Qy	1201	CGATCGCTGAATTTGGCTTGGAA	CAAAATTTGCTATTATTCACCCCA	TGATTTTCCACT 1260
Db	1249	CGATCGCTGAATTTGGCTTGGAA	CAAAATTTGCTATTATTCACCCCA	TGATTTTCCACT 1308
Qy	1261	TTGCCATCCCTAATAAGCTGGACCT	ATCGTCCAACTCTCTGCTCTTTTCC	TCTATAACT 1320
Db	1309	TTGCCATCCCTAATAAGCTGGACCT	ATCGTCCAACTCTCTGCTCTTTTCC	TCTATAACT 1368
Qy	1321	GGGTTACATGGTTTAACTCACTT	TAAATTAACAGGAAATCATGCTTT	TACAGAGCTTGATA 1380
Db	1369	GGGTTACATGGTTTAACTCACTT	TAAATTAACAGGAAATCATGCTTT	TACAGAGCTTGATA 1428
Qy	1381	TCATCTGAAAACCTTTCCAGAACT	CAAGGTTATAGAAATGCTTTATGCTT	TATCCAGTGCTGT 1440
Db	1429	TCATCTGAAAACCTTTCCAGAACT	CAAGGTTATAGAAATGCTTTATGCTT	TATCCAGTGCTGT 1488
Qy	1441	GCATTTGGAGTGTGAGAAATGCCT	ATAAGATTTCTAATCAATGGAAAT	TAAGGTGACAAC 1500
Db	1489	GCATTTGGAGTGTGAGAAATGCCT	ATAAGATTTCTAATCAATGGAAAT	TAAGGTGACAAC 1548

QY	1501	AGCAGTATGAGCGACCTTCTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1566
DB	1549	AGCAGTATGGAACGACCTTCTATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1608
QY	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCATTCAGATGCAG	1620
DB	1609	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCATTCAGATGCAG	1668
QY	1621	TGTTTCACTTCCCAGGCCCCCTTCAAAACCTGTGAAACACCTGCTTGTATGGCTGGCTGATC	1680
DB	1669	TGTTTCACTTCCCAGGCCCCCTTCAAAACCTGTGAAACACCTGCTTGTATGGCTGGCTGATC	1728
QY	1691	AGAAATTGAGTGTGGACCATAGCAGTCTTGCGCATCTTACTTTGTAATGCTTTGGTGTGACTTCA	1740
DB	1729	AGAAATTGAGTGTGGACCATAGCAGTCTTGCGCATCTTACTTTGTAATGCTTTGGTGTGACTTCA	1788
QY	1741	ACAGTTTTAGATCCCCCTCTGTACATTTCCCCCATTTAAACTGTTAAATGGGGTCATCGCA	1800
DB	1789	ACAGTTTTAGATCCCCCTCTGTACATTTCCCCCATTTAAACTGTTAAATGGGGTCATCGCA	1848
QY	1801	GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCCGTGCTGGCTGGTGTGGATGGTTCACCT	1860
DB	1849	GCAGTGAACATGCTCACGGGAGTCTCCAGTGGCCGTGCTGGCTGGTGTGGATGGTTCACCT	1908
QY	1861	TTTGGCAGCTTTGCAACATATGCTGTGGGAGAAATGGGGTTTGGTGTGCCATGTCAATT	1920
DB	1909	TTTGGCAGCTTTGCAACATATGCTGTGGGAGAAATGGGGTTTGGTGTGCCATGTCAATT	1968
QY	1921	GGTTTTTGTCCATTTTTTGCTTCAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGGCCCTG	1980
DB	1969	GGTTTTTGTCCATTTTTTGCTTCAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGGCCCTG	2028
QY	1981	GAGCGTGGGTTCTCTGTGAAATATTCTGCAAAATTTTGAACGAAAGCTCATTTTCTAGC	2040
DB	2029	GAGCGTGGGTTCTCTGTGAAATATTCTGCAAAATTTTGAACGAAAGCTCATTTTCTAGC	2088
QY	2041	CTGAAAGTAATCAATTTTGTCTCTGTGCCCTGCTGGCCTTGACCAATGCCGAGTTCCTCCTG	2100
DB	2089	CTGAAAGTAATCAATTTTGTCTCTGTGCCCTGCTGGCCTTGACCAATGCCGAGTTCCTCCTG	2148
QY	2101	CTGGGTGGCAGCAAGTATGGGCTCTCCCTCTCTGCTGCTGGCTTTGCTTTTGGGAGCCCC	2160
DB	2149	CTGGGTGGCAGCAAGTATGGGCTCTCCCTCTCTGCTGCTGGCTTTGCTTTTGGGAGCCCC	2208
QY	2161	AGCACCATGGCTACATGGTGGCTCTCATCTTCTCAATTCCTTTGCTTTCCTCATGATG	2220
DB	2209	AGCACCATGGCTACATGGTGGCTCTCATCTTCTCAATTCCTTTGCTTTCCTCATGATG	2268
QY	2221	ACCAATGCCCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTGG	2280
DB	2269	ACCAATGCCCTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATATTGG	2328
QY	2281	GACTGTCTATGTGTAAAGACATTTGGCTGTGCTCTTACCAACATGCAATCTTAAACTGC	2340
DB	2329	GACTGTCTATGTGTAAAGACATTTGGCTGTGCTCTTACCAACATGCAATCTTAAACTGC	2388
QY	2341	CCTGTGGCTTTCTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTGGAAGTA	2400
DB	2389	CCTGTGGCTTTCTGTCTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTGGAAGTA	2448
QY	2401	ATTAAGTTTATCTCTTCTGTGGTAGTCCCACTTCTTGCATGTCTCAATCCCCCTTCTCTAC	2460
DB	2449	ATTAAGTTTATCTCTTCTGTGGTAGTCCCACTTCTTGCATGTCTCAATCCCCCTTCTCTAC	2508
QY	2461	ATCTGTATTCAATCTCTCATTTTAAAGAGGATCTGGTGAGCCTTGAGAAAGCAAACTTACGTC	2520
DB	2509	ATCTGTATTCAATCTCTCATTTTAAAGAGGATCTGGTGAGCCTTGAGAAAGCAAACTTACGTC	2568
QY	2521	TGGACAGATCAAAACCAACCAAGCTTGATGTCAATTAATCTTCATGATGTCGAAAAACAG	2580
DB	2569	TGGACAGATCAAAACCAACCAAGCTTGATGTCAATTAATCTTCATGATGTCGAAAAACAG	2628
QY	2581	TCCTGTGACTCAACTCAAGCCTTTGGTAAACCTTTTACCAGCTCTCAGACATCACTTATGACCTG	2640

|||||
441 CTGCTCCGGAATCCCTGCGCCAGTCTCGGCTTCTCGAGGAGTTACGTCTTGGGGAAAC 500
QY
301 GCTCTGACATACATTCCTCAAGGAGCATATCACTGGCCCTTTACAGTCTTAAAGTTCTTTATG 360
Db
501 GCTCTGACATACATTCCTCAAGGAGCATATCACTGGCCCTTTACAGTCTTAAAGTTCTTTATG 560
QY
361 CTCGAGAAATACAGCTTAAGACACGTACCAACGACGATCTGAGAGCTCTGAGAAATTTGCGAAGCCTT 420
Db
561 CTGAGAAATACAGCTTAAGACACGTATACCAACGACGATCTGAGAAATTTGCGAAGCCTT 620
QY
421 CAATCCCTGGCTTGGATGCTTAACCAATCAGTATATGTCGCCCAAGCTGTTTCAGTGGC 480
Db
621 CAATCCCTGGCTTGGATGCTTAACCAATCAGTATATGTCGCCCAAGCTGTTTCAGTGGC 680
QY
481 CTGCAATTCCTGAGGCACTGTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGCTCCAG 540
Db
681 CTGCAATTCCTGAGGCACTGTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGCTCCAG 740
QY
541 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGACAAATACACCAAC 600
Db
741 GCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGACAAATACACCAAC 800
QY
601 ATACCAGACTATGCTTTGGAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 660
Db
801 ATACCAGACTATGCTTTGGAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 860
QY
661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db
861 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 920
QY
721 TTAATTTACAAATACCTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCACCTTAAA 780
Db
921 TTAATTTACAAATACCTTTGATGAATTTCCCACTGCAATTAGGACACTCTCCACCTTAAA 980
QY
781 GAATAGGATTTATAGCAACAATATCAGGTGCGATACCTGAGAAAGCATTTGTAGGCAAC 840
Db
981 GAATAGGATTTATAGCAACAATATCAGGTGCGATACCTGAGAAAGCATTTGTAGGCAAC 1040
QY
841 CCTTCTCTTAATTAACAATTTCTATGACAAATCCCAATCCAAATTTGTTGGGAGATCTGCT 900
Db
1041 CCTTCTCTTAATTAACAATTTCTATGACAAATCCCAATCCCAATTTGTTGGGAGATCTGCT 1100
QY
901 TTTCAACATTTACTGTAACCTAAGAACACTGACTCTGAAATGAGGCTCCACAAATAACTGAA 960
Db
1101 TTTCAACATTTACTGTAACCTAAGAACACTGACTCTGAAATGAGGCTCCACAAATAACTGAA 1160
QY
961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db
1161 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1220
QY
1021 TCATCTCTCCTGAAACCGTCTGGAATCAGTTACCTTAATCTCCAGTGTAGATCTGTCT 1080
Db
1221 TCATCTCTCCTGAAACCGTCTGGAATCAGTTACCTTAATCTCCAGTGTAGATCTGTCT 1280
QY
1081 TACAACTTAATAGAGATTTACCCAGTTTTCAGTCTGCCAAAGCTTTCAGAAATTTGAC 1140
Db
1281 TACAACTTAATAGAGATTTACCCAGTTTTCAGTCTGCCAAAGCTTTCAGAAATTTGAC 1340
QY
1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
Db
1341 CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1400
QY
1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTAATCAACCCCAATGCAATTTCCACT 1260
Db
1401 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTAATCAACCCCAATGCAATTTCCACT 1460
QY
1261 TTGCCATCCCTTAATAAGCTGGAATCTGTCGCAACCTCTGCTGCTTTTCTCTATAACT 1320
Db
1461 TTGCCATCCCTTAATAAGCTGGAATCTGTCGCAACCTCTGCTGCTTTTCTCTATAACT 1520
QY
1321 GGGTTACATGGTTTAACTCACTTTAAAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
Db
|||||

Db
1521 GGGTTACATGGTTTAACTCACTCTTAAATAATTAAACAGGAAATCATGCTTTACAGAGCTTGATA 1580
QY
1381 TCATCTGAAAACTTTTCCAGAACTCAAGGTTTATAGAAATGCTTTATGCTTTACCAAGTGTGT 1440
Db
1581 TCATCTGAAAACTTTTCCAGAACTCAAGGTTTATAGAAATGCTTTATGCTTTACCAAGTGTGT 1640
QY
1441 GCATTTGGAGTGTGTGAGAAATGCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAAAC 1500
Db
1641 GCATTTGGAGTGTGTGAGAAATGCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAAAC 1700
QY
1501 AGCAGTATGGAACGACCTTCTAATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1560
Db
1701 AGCAGTATGGAACGACCTTCTAATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT 1760
QY
1561 GACCTTTGAAGATTTCTGCTTGTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db
1761 GACCTTTGAAGATTTCTGCTTGTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1820
QY
1621 TGTTCACCTTCCCGCAGGCCCTTTCAAAACCTGTGGAACACCTGCTTGATGCTGGCTGATC 1680
Db
1821 TGTTCACCTTCCCGCAGGCCCTTTCAAAACCTGTGGAACACCTGCTTGATGCTGGCTGATC 1880
QY
1681 AGAATTTGGAGTGTGGAACCATPAGCAGTTCTGGCACTTACTTTGTAATGCTTTTGGTGAATCA 1740
Db
1881 AGAATTTGGAGTGTGGAACCATPAGCAGTTCTGGCACTTACTTTGTAATGCTTTTGGTGAATCA 1940
QY
1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAATACTGTTAATTTGGGTCATCGCA 1800
Db
1941 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAATACTGTTAATTTGGGTCATCGCA 2000
QY
1801 GCAGTGAAACATGCTCAACGGAGTCTCCAGTCCGCTGCTGGCTGGTGTGGAATGCTTCACT 1860
Db
2001 GCAGTGAAACATGCTCAACGGAGTCTCCAGTCCGCTGCTGGCTGGTGTGGAATGCTTCACT 2060
QY
1861 TTTGGCAGCTTTGCAACGACATGCTGCTGGTGGGAGAAATGGGGTGTGGTCCCATGTCATT 1920
Db
2061 TTTGGCAGCTTTGCAACGACATGCTGCTGGTGGGAGAAATGGGGTGTGGTCCCATGTCATT 2120
QY
1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTCTGCTTACTCTGCGACGCCCTG 1980
Db
2121 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTCTGCTTACTCTGCGACGCCCTG 2180
QY
1981 GACGTGGGGTCTCTGTGAAATATTCTGCAAAATTTTGAACGAAAGCTCCATTTCTTAGC 2040
Db
2181 GACGTGGGGTCTCTGTGAAATATTCTGCAAAATTTTGAACGAAAGCTCCATTTCTTAGC 2240
QY
2041 CTGAAAGTAAATCATTTTGTCTCTGCTGCTGCTGGCTGTGACCAATGCGCGCAGTTCCCTG 2100
Db
2241 CTGAAAGTAAATCATTTTGTCTCTGCTGCTGCTGGCTGTGACCAATGCGCGCAGTTCCCTG 2300
QY
2101 CTGGGTGGCAGCAAGTATGCGCCTCCCTCTCTGCTGCTGCTTTTGGGGAGCC 2160
Db
2301 CTGGGTGGCAGCAAGTATGCGCCTCCCTCTCTGCTGCTGCTTTTGGGGAGCC 2360
QY
2161 AGCACCATGCGCTACATGCTGCTCATCTGCTCAAATTTCCCTTTTGGCTTCTCATGATG 2220
Db
2361 AGCACCATGCGCTACATGCTGCTCATCTGCTCAAATTTCCCTTTTGGCTTCTCATGATG 2420
QY
2221 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATTTGG 2280
Db
2421 ACCATTTGCCCTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATTTGG 2480
QY
2281 GACTGCTTATGTTGTAACACATTTGCCCTGCTCTTCCACCAACCTGATCTCTAAACCTGC 2340
Db
2481 GACTGCTTATGTTGTAACACATTTGCCCTGCTCTTCCACCAACCTGATCTCTAAACCTGC 2540
QY
2341 CTTGTGGCTTTCTTGTCTCTCTCTTCTTAAATAAACCTTACATTTATFAGTCCCTGAAGTA 2400
Db
2541 CTTGTGGCTTTCTTGTCTCTCTCTTCTTAAATAAACCTTACATTTATFAGTCCCTGAAGTA 2600
QY
2401 ATTAAGTTTATCTTCTGTTGGTGGTGGTCCACTTCTGCTGATGCTCAATCCCTTCTCTAC 2460
Db
2601 ATTAAGTTTATCTTCTGTTGGTGGTGGTCCACTTCTGCTGATGCTCTCAATCCCTTCTCTAC 2660

2281 GACTGCTCTATGGTAAACACATATGGCCCTGTGGCTCTTCCACCAACTGCATCTAAACTGC 2340
2281 GACTGCTCTATGGTAAACACATATGGCCCTGTGGCTCTTCCACCAACTGCATCTAAACTGC 2340
2341 CCGTGGGCTTCTGTGCT 2400
2341 CCGTGGGCTTCTGTGCT 2400
2401 ATTAAGTTTATCTCTGTGGTAGTCCCACTTCTCTGCAATGCTCAATCCCTCTCTCTAC 2460
2401 ATTAAGTTTATCTCTGTGGTAGTCCCACTTCTCTGCAATGCTCAATCCCTCTCTCTAC 2460
2461 ATCTGTTTCAATCTCTCACTTTAAGGAGGATCTGGTAGGCTCAGAAAGCAACTACGTC 2520
2461 ATCTGTTTCAATCTCTCACTTTAAGGAGGATCTGGTAGGCTCAGAAAGCAACTACGTC 2520
2521 TGGACAAGATCAAAACACCAAGCTTGTATGCTCAATTTAACTCTGTATGATGTGCAAAAACAG 2580
2521 TGGACAAGATCAAAACACCAAGCTTGTATGCTCAATTTAACTCTGTATGATGTGCAAAAACAG 2580
2581 TCCTGTGACTCAACTCAAGCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
2581 TCCTGTGACTCAACTCAAGCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
2641 CTTCCAGTTCGGTGCCATCAGCTTATCCAGTCACTGAGAGCTGCCATCTTTCTCTCT 2700
2641 CTTCCAGTTCGGTGCCATCAGCTTATCCAGTCACTGAGAGCTGCCATCTTTCTCTCT 2700
2701 GTGGCAATTTGCCATGCTC 2721
2701 GTGGCAATTTGCCATGCTC 2721

RESULT 12
US-10-251-385-277
; Sequence 277, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251,385
; CURRENT FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 277
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-277

Query Match 96.3%; Score 2622; DB 15; Length 2724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2722; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGGACACTCCGGCTCGGTGTGCT 60
1 ATGGACACTCCGGCTCGGTGTGCT 60
61 GGGGGCAGCTCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCACACACTGTCTATGTC 120
61 GGGGGCAGCTCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCACACACTGTCTATGTC 120
121 GAGCCGCGAGGAGATGTGCTCAGGGTGGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 180
121 GAGCCGCGAGGAGATGTGCTCAGGGTGGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 180

QY 181 CTTCCAACTCAGCGTCTTCACTCCTCACTAGACCTCAGTATGAAACAACATCAGTCAG 240
DB 181 CTTCCAACTCAGCGTCTTCACTCCTCACTAGACCTCAGTATGAAACAACATCAGTCAG 240
QY 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTTCTCTGGAGGAGTTACGCTTCTCGGGAAC 300
DB 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTTCTCTGGAGGAGTTACGCTTCTCGGGAAC 300
QY 301 GCTCTGACATACATTTCCAAAGGAGCATTCATCTGGCCCTTTACAGTCTTAAAGTTCTTATG 360
DB 301 GCTCTGACATACATTTCCAAAGGAGCATTCATCTGGCCCTTTACAGTCTTAAAGTTCTTATG 360
QY 361 CTGCAAGATATCAGCTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCTT 420
DB 361 CTGCAAGATATCAGCTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCTT 420
QY 421 CAATCCCTGCGTCTGGATGTCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGCG 480
DB 421 CAATCCCTGCGTCTGGATGTCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGCG 480
QY 481 CTGCAATCCCTGAGGCACTGTGTGCTGATGACAAATGCGTTAAACAGAAATTTCCCGTCCAG 540
DB 481 CTGCAATCCCTGAGGCACTGTGTGCTGATGACAAATGCGTTAAACAGAAATTTCCCGTCCAG 540
QY 541 GCTTTTAGAAGTTTATCGGCATTCGAGCCATGCAAGCCATGACCTTGGCCCTTGAACAAATACACCA 600
DB 541 GCTTTTAGAAGTTTATCGGCATTCGAGCCATGCAAGCCATGACCTTGGCCCTTGAACAAATACACCA 600
QY 601 ATACAGACTATGCTTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 660
DB 601 ATACAGACTATGCTTTTGGAAACCTCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 660
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTGTATGGGCTCCACAGCCTAGAGACTTTAGAT 720
DB 661 AGAATCCACTCCCTGGGAAAGAAATGCTTGTATGGGCTCCACAGCCTAGAGACTTTAGAT 720
QY 721 TTAATTAACAATTAACCTTGTATGAAATTCGCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
DB 721 TTAATTAACAATTAACCTTGTATGAAATTCGCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
QY 781 GAACTAGGATTTTCATAGCAAAATATCAGGTGCGATACCTCGAAGAAAGCAATTTGTAGGCAAC 840
DB 781 GAACTAGGATTTTCATAGCAAAATATCAGGTGCGATACCTCGAAGAAAGCAATTTGTAGGCAAC 840
QY 841 CCTCTCTTATTAACAATCAATTTCTATGACAAATCCCATCCCAATTTTGTGGGAGATCTGCT 900
DB 841 CCTCTCTTATTAACAATCAATTTCTATGACAAATCCCATCCCAATTTTGTGGGAGATCTGCT 900
QY 901 TTTTCAACATTTTACCTGAACTAAGACACTGACTCTGAATGGTGGCTCACAATAAATCTGAA 960
DB 901 TTTTCAACATTTTACCTGAACTAAGACACTGACTCTGAATGGTGGCTCACAATAAATCTGAA 960
QY 961 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
DB 961 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
QY 1021 TCATCTCTTCTCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCCAAAGTGTAGATCTGCT 1080
DB 1021 TCATCTCTTCTCTCAAAACCGTCTGCAATCAGTTTACCTTAATCTCCAAAGTGTAGATCTGCT 1080
QY 1081 TACAACTTATTAAGAAATTTACCCAGTCTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1140
DB 1081 TACAACTTATTAAGAAATTTACCCAGTCTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1140
QY 1141 CTAAGACATTAATGAAATCTACGAAATTAAGTGTGACATTTTCCAGCAGTGTCTTAGCCTC 1200
DB 1141 CTAAGACATTAATGAAATCTACGAAATTAAGTGTGACATTTTCCAGCAGTGTCTTAGCCTC 1200
QY 1201 CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTTCACCCCAATGCAATTTTCCACT 1260
DB 1201 CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTTCACCCCAATGCAATTTTCCACT 1260

QY 1261 TTGCGATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTCTTTCTTAACCT 1320
DB 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCTCTGCTCTTTCTTAACCT 1320
QY 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1380
DB 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTCTGT 1440
DB 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTCTGT 1440
QY 1441 GCATTTGGAGTGTGAGATGCTTATAGATTTCTAATCAATGGAATGAAGTGAAC 1500
DB 1441 GCATTTGGAGTGTGAGATGCTTATAGATTTCTAATCAATGGAATGAAGTGAAC 1500
QY 1501 AGCAGTATGACGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAC 1560
DB 1501 AGCAGTATGACGACCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAC 1560
QY 1561 GACCTTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
DB 1561 GACCTTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
QY 1621 TGTTCACCTTCCCAGGCGCCCTTCAAAACCTGTGAACACCTGCTGATGCTGGCTGATC 1680
DB 1621 TGTTCACCTTCCCAGGCGCCCTTCAAAACCTGTGAACACCTGCTGATGCTGGCTGATC 1680
QY 1681 AGAATTTGGAGTGTGAGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTGGTGACTTCA 1740
DB 1681 AGAATTTGGAGTGTGAGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTGGTGACTTCA 1740
QY 1741 ACAGTTTTCAGATFCCCTCTGTACATTTCCCACTTAAATGTTAAATGGGGTCATCGCA 1800
DB 1741 ACAGTTTTCAGATFCCCTCTGTACATTTCCCACTTAAATGTTAAATGGGGTCATCGCA 1800
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCAGTCCGCTGCTGGCTGGTGTGATGATGTTCACT 1860
DB 1801 GCAGTGAACATGCTCAAGGAGTCTCAGTCCGCTGCTGGCTGGTGTGATGATGTTCACT 1860
QY 1861 TTTGGCAGCTTTGCAGACATGCTGCTGCTGGGAGAAATGGGTTGGTGGCATGTCATT 1920
DB 1861 TTTGGCAGCTTTGCAGACATGCTGCTGCTGGGAGAAATGGGTTGGTGGCATGTCATT 1920
QY 1921 GGTGTTTTGTCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
DB 1921 GGTGTTTTGTCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
QY 1981 GAGCGTGGGTTCTCTGTAATAATTTCTGAAATTTTGAACGAAAGCTCAATTTTCTAGC 2040
DB 1981 GAGCGTGGGTTCTCTGTAATAATTTCTGAAATTTTGAACGAAAGCTCAATTTTCTAGC 2040
QY 2041 CTGAAAGTAAATCAATTTGCTCTGTCGCCCTGCTGGCTTGACCAATGSCCAGTTCCCTG 2100
DB 2041 CTGAAAGTAAATCAATTTGCTCTGTCGCCCTGCTGGCTTGACCAATGSCCAGTTCCCTG 2100
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGGGAGGCC 2160
DB 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGGGAGGCC 2160
QY 2161 AGCACATGGGTATCATGCTGCTCTCATCTTGTCAATTCCTTTGCTTCTCATGATG 2220
DB 2161 AGCACATGGGTATCATGCTGCTCTCATCTTGTCAATTCCTTTGCTTCTCATGATG 2220
QY 2221 ACCATTTGCTTACACCAAGCTCTACTGCAATTTGGACAGGGAGCTGGAGATATTTGG 2280
DB 2221 ACCATTTGCTTACACCAAGCTCTACTGCAATTTGGACAGGGAGCTGGAGATATTTGG 2280
QY 2281 GACTGCTCTATGTAACCAACATTTGCTGCTCTTCCAACTGCACTCTTAACTGC 2340
DB 2281 GACTGCTCTATGTAACCAACATTTGCTGCTCTTCCAACTGCACTCTTAACTGC 2340
QY 2341 CCTGNGGCTTTCTTGTCTCTCTCTTTAATAAACCTTTACATTTATCAGTCTGAAGTA 2400

DB 2341 CCTGNGGCTTTCTTGTCTCTCTCTTTAATAAACCTTTACATTTATCAGTCTGAAGTA 2400
QY 2401 ATTAAAGTTTATCTCTGCTGCTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
DB 2401 ATTAAAGTTTATCTCTGCTGCTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTAC 2460
QY 2461 ATCTGTTTCAATCTCTCACTTTTAAAGGAGGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2520
DB 2461 ATCTGTTTCAATCTCTCACTTTTAAAGGAGGATCTGGTGGAGCTGAGAAAGCAACCTACGTC 2520
QY 2521 TGGACAAGATCAAAACACCAAGCTTGTCAATTTAATCTGTATGATGTGCAAAACAG 2580
DB 2521 TGGACAAGATCAAAACACCAAGCTTGTCAATTTAATCTGTATGATGTGCAAAACAG 2580
QY 2581 TCTGTGACTCAACTCAAGCTTGTGAACCTTTTACCAGCTCCAGCATCACTTATGACCTG 2640
DB 2581 TCTGTGACTCAACTCAAGCTTGTGAACCTTTTACCAGCTCCAGCATCACTTATGACCTG 2640
QY 2641 CTCCAGTTCGCTGCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCT 2700
DB 2641 CTCCAGTTCGCTGCCATCACAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700
QY 2701 GTGGCATTTGTCCCATGTCTCTAA 2724
DB 2701 GTGGCATTTGTCCCATGTCTCTAA 2724

RESULT 13

US-10-295-027-483
; Sequence 483, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; TITLE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 483
; LENGTH: 2651

; TYPE: DNA									
; ORGANISM: Homo sapiens									
US-10-295-027-483									
Query Match 68.4%; Score 1864; DB 17; Length 2651;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1864; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	857	TACATTTCTATGACAAATCCATCCAAATTTGTTGGAGAGTCTGCTTTTCAACATTTACCTG	916		1745	CGGAGTCTCCAGTGCCTGCTGGCTGCTGGATGCGTTCACTTTTGGCAGCTTTGCAC	1804		
Db	785	TACATTTCTATGACAAATCCATCCAAATTTGTTGGAGAGTCTGCTTTTCAACATTTACCTG	844		1877	GACATGGTCCCTGGTGGGAGAAATGGGGTTGGTGGCATGTCAATGGTTTTTGTTCATTT	1936		
QY	917	ACTAAGACACTGACTCTGAATGTGCTGCTCAAAATTAACAAATTTCTGATTTAACTG	976		1805	GACATGGTGCCTGGTGGGAGAAATGGGGTTGGTGGCATGTCAATGGTTTTTGTTCATTT	1864		
Db	845	AACTAAGAACACTGACTCTGAATGGTGCCTCAAAATTAACAAATTTCTGATTTAACTG	904		1937	TTGCTTCAGAAATCATCTGTTTTTCCCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTCTCTG	1996		
QY	977	GAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGCAGACATCTCATCTCTCTCTCAAA	1036		1865	TTGCTTCAGAAATCATCTGTTTTTCCCTGCTTACTCTGGCAGCCCTGGAGCGTGGGTCTCTG	1924		
Db	905	GAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGCAGACATCTCATCTCTCTCTCAAA	964		1997	TGAATAATTTCTGCAAAATTTGAAAACGAAAGCTCCATTTTTCTAGCCCTGAAAAGTAATCA	2056		
QY	1037	CCGCTGCAATCAGTTACCTAATCTCCAAAGTGTAGATCTGCTTACAACTTATAGAG	1096		1925	TGAATAATTTCTGCAAAATTTGAAAACGAAAGCTCCATTTTTCTAGCCCTGAAAAGTAATCA	1984		
Db	965	CCGCTGCAATCAGTTACCTAATCTCCAAAGTGTAGATCTGCTTACAACTTATAGAG	1024		2057	TGCTCTGTCCTGCTGCTGCTGCTGACCATGGCGCGATGTTCCCTGCTGGGTGGCAGCAAGT	2116		
QY	1097	ATTTACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAATTAACCTAAGACATAATGAAA	1156		1985	TGCTCTGTCCTGCTGCTGCTGCTGACCATGGCGCGATGTTCCCTGCTGGGTGGCAGCAAGT	2044		
Db	1025	ATTTACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAATTAACCTAAGACATAATGAAA	1084		2117	ATGGCGCTCCCTCTCTGCTGCTGCTGCTTTGGGGAGCCAGCACCATGGGCTACA	2176		
QY	1157	TCTAGGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTCGATCGCTGAAATTTGG	1216		2045	ATGGCGCTCCCTCTCTGCTGCTGCTTTGGCCCTTTGGGGAGCCAGCACCATGGGCTACA	2104		
Db	1085	TCTAGGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTCGATCGCTGAAATTTGG	1144		2177	TGGTGCCTCTCATCTTTGCTCAATTTCCCTTCTCTCATGATGACCATTTGCCTACACA	2236		
QY	1217	CTTGGAACAAATATGCTATATATACCCCAATGCAATTTTCCATCTTGCCATCCCTAATAA	1276		2105	TGGTGCCTCTCATCTTTGCTCAATTTCCCTTCTCTCATGATGACCATTTGCCTACACA	2164		
Db	1145	CTTGGAACAAATATGCTATATATACCCCAATGCAATTTTCCATCTTGCCATCCCTAATAA	1204		2237	AGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAAATATTTGGGACGTCTATGGTAA	2296		
QY	1277	AGCTGACCTATGCTCCACCTCTGCTGCTTTTCCCTAATACCTGGTTACATGGTTTAA	1336		2165	AGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAAATATTTGGGACGTCTATGGTAA	2224		
Db	1205	AGCTGACCTATGCTCCACCTCTGCTGCTTTTCCCTAATAACTGGTTACATGGTTTAA	1264		2297	AACATATGGCCCTGTTGCTCTTCCAACTGCATCTAAACTGCCCTGGCTTTCTGT	2356		
QY	1337	CTCATTAAAAATTAACAGGAAATCATGCTTACAGAGCTTGATATCATCTGAAAACTTTC	1396		2225	AACATATGGCCCTGTTGCTCTTCCAACTGCATCTAAACTGCCCTGGCTTTCTGT	2284		
Db	1265	CTCATTAAAAATTAACAGGAAATCATGCTTACAGAGCTTGATATCATCTGAAAACTTTC	1324		2357	CCTTCTCCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAAGTAAATTAAGTTTATCCTTC	2416		
QY	1397	CAGAACTCAAGTTATAGAAATGCTTATGCTTTAACAAGTCTGTGATTTGAGTGTGTG	1456		2285	CCTTCTCCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAAGTAAATTAAGTTTATCCTTC	2344		
Db	1325	CAGAACTCAAGTTATAGAAATGCTTATGCTTTACCAAGTCTGTGCTATTTGGAGTGTGTG	1384		2417	TGGTGGTAGTCCACTTCTGATGTCTCAATCCCTTCTCTACATCTTGTTCATCTCTC	2476		
QY	1457	AGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGACAAACAGCAGTATGGACGACC	1516		2345	TGGTGGTAGTCCACTTCTGATGTCTCAATCCCTTCTCTACATCTTGTTCATCTCTC	2404		
Db	1385	AGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGACAAACAGCAGTATGGACGACC	1444		2477	ACTTTAAAGAGGATCTGCTGAGCCCTGAGAAAGCAAACTTACCTCTGGCAAGATCAAAAC	2536		
QY	1517	TTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGCTGACCTTGAAGATTTCC	1576		2405	ACTTTAAAGAGGATCTGCTGAGCCCTGAGAAAGCAAACTTACCTCTGGCAAGATCAAAAC	2464		
Db	1445	TTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGCTGACCTTGAAGATTTCC	1504		2537	ACCCAAGCTTGATGTCAATTTAACTCTGATGATGTCGAAAAACAGTCTCTGTGACTCAACTC	2596		
QY	1577	TGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGAAGTGTTCACCTTCCCGAG	1636		2465	ACCCAAGCTTGATGTCAATTTAACTCTGATGATGTCGAAAAACAGTCTCTGTGACTCAACTC	2524		
Db	1505	TGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGAAGTGTTCACCTTCCCGAG	1564		2597	AGCCCTTGGTAAACCTTTACAGCTCCAGCATCACTTATGACCTGCCAGTTCGGTGC	2656		
QY	1637	GCCTTCAAAACCTGTGGAACACCTGCTGATGGCTGGCTGATCAGAAATTTGAGTGTGA	1696		2525	AGCCCTTGGTAAACCTTTTACAGCTCCAGCATCACTTATGACCTGCCCTCCAGTTCGGTGC	2584		
Db	1565	GCCTTCAAAACCTGTGGAACACCTGCTGATGGCTGGCTGATCAGAAATTTGAGTGTGA	1624		2657	CATCACCAGCTTATCCAGTGAAGTGCATCTTTTCCCTCTGTGGCAATTTGTCCCAT	2716		
QY	1697	CCATAGCAGTTCTGGCACTTACTTTGTAATGCTTTGGTGAACCTTCAACAGTTTTTTCAGATCCC	1756		2585	CATCACCAGCTTATCCAGTGAAGTGCATCTTTTCCCTCTGTGGCAATTTGTCCCAT	2776		
Db	1625	CCATAGCAGTTCTGGCACTTACTTTGTAATGCTTTGGTGAACCTTCAACAGTTTTTTCAGATCCC	1684		2717	GTCT 2720			
QY	1757	CTCTGTACATTTCCCTCCCTAACTGTTAAATTTGGGTGATTCGACAGTGAACATGCTCA	1816		2645	GTCT 2648			
Db	1685	CTCTGTACATTTCCCTCCCTAACTGTTAAATTTGGGTGATTCGACAGTGAACATGCTCA	1744						
QY	1817	CGGAGTCTCCAGTGCCTGCTGGCTGGTGTGATGCGTTCACTTTTGGCAGCTTTGCAC	1876						

RESULT 14
US-10-295-027-1113
; Sequence 1113, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziiz, Natsaha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.


```

: APPLICANT: Glyne, Richard
: APPLICANT: Hevezi, Peter A.
: APPLICANT: Mack, David H.
: APPLICANT: Murray, Richard
: APPLICANT: Watson, Susan R.
: APPLICANT: Eos Biotechnology, Inc.
: TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
: FILE REFERENCE: 018501-012500US
: CURRENT APPLICATION NUMBER: US/10/295,027
: CURRENT FILING DATE: 2002-11-13
: PRIOR APPLICATION NUMBER: US 09/663,733
: PRIOR FILING DATE: 2000-09-15
: PRIOR APPLICATION NUMBER: US 60/350,666
: PRIOR FILING DATE: 2001-11-13
: PRIOR APPLICATION NUMBER: US 60/335,394
: PRIOR FILING DATE: 2001-11-15
: PRIOR APPLICATION NUMBER: US 60/332,464
: PRIOR FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: US 60/334,393
: PRIOR FILING DATE: 2001-11-29
: PRIOR APPLICATION NUMBER: US 60/340,376
: PRIOR FILING DATE: 2001-12-14
: PRIOR APPLICATION NUMBER: US 60/347,211
: PRIOR FILING DATE: 2002-01-08
: PRIOR APPLICATION NUMBER: US 60/347,349
: PRIOR FILING DATE: 2002-01-10
: PRIOR APPLICATION NUMBER: US 60/355,250
: PRIOR FILING DATE: 2002-02-08
: PRIOR APPLICATION NUMBER: US 60/356,714
: PRIOR FILING DATE: 2002-02-13
: Remaining Prior Application data removed - See File Wrapper or PALM.
: NUMBER OF SEQ ID NOS: 1386
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1113
: LENGTH: 2651
: TYPE: DNA
: ORGANISM: Homo sapiens
: US-10-295-027-1113

Query Match          69.4%; Score 1864; DB 17; Length 2651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      857  TACATTTCTATGCAATCCCATCCAAATTTGTTGGAGATCTGCTTTTCAACATTTACCTG 916
      |||||
Db      785  TACATTTCTATGCAATCCCATCCAAATTTGTTGGAGATCTGCTTTTCAACATTTACCTG 844

Qy      917  AACTAAGAACACTGACTCTGAATGGTGCCTCAACAAATACATGAATTTCTCGATTTAACTG 976
      |||||
Db      845  AACTAAGAACACTGACTCTGAATGGTGCCTCAACAAATACATGAATTTCTCGATTTAACTG 904

Qy      977  GAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATCTCATCTCTCTCTCAA 1036
      |||||
Db      905  GAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATCTCATCTCTCTCTCAA 964

Qy      1037  CCGTCTGCAATCAGTTACCTAACTCTCAAGTGCTAGATCTGTCTTACAACTATTAGAAG 1096
      |||||
Db      965  CCGTCTGCAATCAGTTACCTAACTCTCAAGTGCTAGATCTGTCTTACAACTATTAGAAG 1024

Qy      1097  ATTTACCCAGTTTTTCAGTCTGCGCAAAAGCTTCAGAAAAATGACCTTAAGACATAATGAAA 1156
      |||||
Db      1025  ATTTACCCAGTTTTTCAGTCTGCGCAAAAGCTTCAGAAAAATGACCTTAAGACATAATGAAA 1084

Qy      1157  TCTACGAAATTAAGTTGACACTTTTCAGCAGAGTTGCTTAGCCTCCGATCGCTGAATTTGG 1216
      |||||
Db      1085  TCTACGAAATTAAGTTGACACTTTTCAGCAGAGTTGCTTAGCCTCCGATCGCTGAATTTGG 1144

Qy      1217  CTTGGAAACAAATTTGCTATTATTTCACCCCAATGCATTTTCCACTTTTCCCATCCTCAATAA 1276
      |||||
Db      1145  CTTGGAAACAAATTTGCTATTATTTCACCCCAATGCATTTTTCACCTTTGCCATCCTCAATAA 1204

Qy      1277  AGCTGGACCTATCGTCCAAACCTCTCTGCTCTTTTCTTAACTATGCTGGGTATACATGGTTTAA 1336
      |||||

```


Db 2285 CCTTCTCCTCTTTTAATAAACCTTACATTTATATCAGTCTGGAAGTAATTAAGTTTATCCTTC 2344
Qy 2417 TGGTGGAGTCCCACTTCTCCTGATGTCTCAATFCCCTTCTCTA CATCTGTTGTTCAATCCTC 2476
Db 2345 TGGTGGTAGTCCCACTTCTCCTGATGTCTCAATFCCCTTCTCTACATCTGTTTCAATCCTC 2404
Qy 2477 ACTTTAAGGAGTCTGGTGGAGCTGGAAGCAAACTAGCTCTGGACAGATCAAAAC 2536
Db 2405 ACTTTAAGGAGGATCTGGTGGAGCTGGAAGCAAACTAGCTCTGGACAGATCAAAAC 2464
Qy 2537 ACCCAAGCTTGATGTCAATTAACCTCTGATGATGTGCAAAACAGTCTCTGTGACTCAACTC 2596
Db 2465 ACCCAAGCTTGATGTCAATTAACCTCTGATGATGTGCAAAACAGTCTCTGTGACTCAACTC 2524
Qy 2597 AAGCTTGGTAACTTTACAGCTCCAGCATCACTTATGACCTGCTCCAGTTCCGTGC 2656
Db 2525 AAGCTTGGTAACTTTACAGCTCCAGCATCACTTATGACCTGCTCCAGTTCCGTGC 2584
Qy 2657 CATCACCAGCTTATCCAGTGAAGTGAAGCTGCCATCTTCTCTGTGGCATTTGTCCTCAT 2716
Db 2585 CATCACCAGCTTATCCAGTGAAGTGAAGCTGCCATCTTCTCTGTGGCATTTGTCCTCAT 2644
Qy 2717 GTCT 2720
Db 2645 GTCT 2648

RESULT 15

US-10-173-999-27
; Sequence 27, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Ovarian
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002420US
; CURRENT APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 27
; LENGTH: 2651
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-173-999-27

Query Match 68.4%; Score 1864; DB 17; Length 2651;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1864; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 857 TACATTTCTATGACATCCCAATCTTGTGGAGATCTGCTTTTCAACATTTACTG 916
Db 785 TACATTTCTATGACAAATCCCAATCTTGTGGAGATCTGCTTTTCAACATTTACTG 844
Qy 917 AACTAAGAACACTGACTCTGAATGTGCTCACAATAAATGAATTTCTCATTTAACTG 976
Db 845 AACTAAGAACACTGACTCTGAATGTGCTCACAATAAATGAATTTCTCATTTAACTG 904
Qy 977 GAATGTCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATCTCATCTTCTCTCAA 1036
Db 905 GAATGTCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATCTCATCTTCTCTCAA 964

Qy 1037 CCGTCTGCAATCAGTTACCTAAATCTCCAAGTGTAGATCTGTCTTACAACTATTAGAAG 1096
Db 965 CCGTCTGCAATCAGTTACCTAAATCTCCAAGTGTAGATCTGTCTTACAACTATTAGAAG 1024
Qy 1097 ATTTACCAGTCTTTCAGTCTGCCAAAGCTTTCAGAAAAATTTGACCTAAGACATATGAAA 1156
Db 1025 ATTTACCAGTCTTTCAGTCTGCCAAAGCTTTCAGAAAAATTTGACCTAAGACATATGAAA 1084
Qy 1157 TCTACGAAATTAAGTTTGACACTTTTCAGAGTGTCTTAGCCTCCGATCGTGAATTTGG 1216
Db 1085 TCTACGAAATTAAGTTTGACACTTTTCAGAGTGTCTTAGCCTCCGATCGTGAATTTGG 1144
Qy 1217 CTTTGGACAAATTTGCTATTATTACACCCCAATGCTTTTCCACTTTCCCAATCCCTAATAA 1276
Db 1145 CTTTGGACAAATTTGCTATTATTACACCCCAATGCTTTTCCACTTTCCCAATCCCTAATAA 1204
Qy 1277 AGCTGGACCTATCGTCCAACTCTCTGTGCTTTTCCCTATAAATCTGGGTTCATGTTTAA 1336
Db 1205 AGCTGGACCTATCGTCCAACTCTCTGTGCTTTTCCCTATAAATCTGGGTTCATGTTTAA 1264
Qy 1337 CTCACTTAAATTAAGGAAATCATGCTTACAGGCTTGTATCATCTGAAATCTTC 1396
Db 1265 CTCACTTAAATTAAGGAAATCATGCTTACAGGCTTGTATCATCTGAAATCTTC 1324
Qy 1397 CAGAACTCAAGGTTATAGAAATGCTTTATCCTTACAGTGTGTGCTTTGGAGTGTGTG 1456
Db 1325 CAGAACTCAAGGTTATAGAAATGCTTTATCCTTACAGTGTGTGCTTTGGAGTGTGTG 1384
Qy 1457 AGAATGCTTATAAGATTTCTAATCAATGGAATTAAGGTGACAAAGAGTATGAGCAACC 1516
Db 1385 AGAATGCTTATAAGATTTCTAATCAATGGAATTAAGGTGACAAAGAGTATGAGCAACC 1444
Qy 1517 TTCAATAAGAAAGTCTGGAAATGTTTTCAGGCTCAAGATGAACGCTTGAAGATTTCC 1576
Db 1445 TTCAATAAGAAAGTCTGGAAATGTTTTCAGGCTCAAGATGAACGCTTGAAGATTTCC 1504
Qy 1577 TGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGTGCTTCACTTCCCGAG 1636
Db 1505 TGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGTGCTTCACTTCCCGAG 1564
Qy 1637 GCCCTTTCAAAACCTGTGAACACCTGCTTGATGGCTGGCTGATCAGAAATTTGGAGTGA 1696
Db 1565 GCCCTTTCAAAACCTGTGAACACCTGCTTGATGGCTGGCTGATCAGAAATTTGGAGTGA 1624
Qy 1697 CCATAGCAGTCTGGGCACTTACTTGTGATGCTTGTGCTTCAACAGTTTTCAGATCCC 1756
Db 1625 CCATAGCAGTCTGGGCACTTACTTGTGATGCTTGTGCTTCAACAGTTTTCAGATCCC 1684
Qy 1757 CTCTGTACATTTCCCTCATTTAACTGTTAATTTGGGCTCATCGCAGCAGTGAACATGCTCA 1816
Db 1685 CTCTGTACATTTCCCTCATTTAACTGTTAATTTGGGCTCATCGCAGCAGTGAACATGCTCA 1744
Qy 1817 CGGGAGTCTCCAGTGGCTGCTGGCTGTGATGCTTCACTTTTGGAGCAGCTTTGAC 1876
Db 1745 CGGGAGTCTCCAGTGGCTGCTGGCTGTGATGCTTCACTTTTGGAGCAGCTTTGAC 1804
Qy 1877 GACATGCTGCTGGTGGAGAAATGGGGTGTGGTGGCAATGCTTGTGTTTGTGCCATTT 1936
Db 1805 GACATGCTGCTGGTGGAGAAATGGGGTGTGGTGGCAATGCTTGTGTTTGTGCCATTT 1864
Qy 1937 TTGCTTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTGGAGCGTGGTCTCTG 1996
Db 1865 TTGCTTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTGGAGCGTGGTCTCTG 1924
Qy 1997 TGAATAATTTCTGCAAAATTTGAAACGAAAGCTCAATTTCTAGCTGAAAGTAAATCATTT 2056
Db 1925 TGAATAATTTCTGCAAAATTTGAAACGAAAGCTCAATTTCTAGCTGAAAGTAAATCATTT 1984
Qy 2057 TGCTCTGCTGCTGCTGGCTTGAACATGCGCGAGTTTCCCTGCTGCTGGTGGCAGCAAGT 2116
Db 1985 TGCTCTGCTGCTGCTGGCTTGAACATGCGCGAGTTTCCCTGCTGCTGGTGGCAGCAAGT 2044
Qy 2117 ATGGCGCTCTCCCTCTCTGCTGCTTTTGGCTTTTGGGAGAGCCAGCAGCATGGGTACA 2176

Db	2045		ATGCGCCTCCCTCTCTGCGCTTGTGCTTTGGGAGCCGACCATGGGCTACA	2104
Qy	2177	TGTCGCTCTCATCTTGCTCAATTCCCTTTGCTTTCCTCATGATGACCAATTGCCCTACACCA	2236	
Db	2105	TGTCGCTCTCATCTTGCTCAATTCCCTTTGCTTTCCTCATGATGACCAATTGCCCTACACCA	2164	
Qy	2237	AGCTCTACTGGAATTTGGCAAGGGAGACCTGGAGAAATATTTGGGACTGCTCTATGGTAA	2296	
Db	2165	AGCTCTACTGGAATTTGGCAAGGGAGACCTGGAGAAATATTTGGGACTGCTCTATGGTAA	2224	
Qy	2297	AACACATTGGCCCTTGCTCTTTACCAACTGCAATCCCTAAACTGCCCTGGGCTTTCTTGT	2356	
Db	2225	AACACATTGGCCCTTGCTCTTTACCAACTGCAATCCCTAAACTGCCCTGGGCTTTCTTGT	2284	
Qy	2357	CTTCTCTCTTTTAATAAAAGCTTACATTTATCAGTCTCTGAAGTAATTAAGTTTATCGTTTC	2416	
Db	2285	CTTCTCTCTTTTAATAAAAGCTTACATTTATCAGTCTCTGAAGTAATTAAGTTTATCGTTTC	2344	
Qy	2417	TGTCGTGAGTCCCACTTTCCTGCATGTCCTCAATCCCTTCTCTACATCTTGCTCAATCCCTC	2476	
Db	2345	TGTCGTGAGTCCCACTTTCCTGCATGTCCTCAATCCCTTCTCTACATCTTGCTCAATCCCTC	2404	
Qy	2477	ACTTTAAGGAGGATCTGGTGAGCGCTGAGAAAGCAAACTTACGTCTGGAAGAATCAAAC	2536	
Db	2405	ACTTTAAGGAGGATCTGGTGAGCGCTGAGAAAGCAAACTTACGTCTGGAAGAATCAAAC	2464	
Qy	2537	ACCCAGCTTGATGTCATTAATTAACCTTCATGATGTCGAAAAACAGTCTCTGTGACTCAACTC	2596	
Db	2465	ACCCAGCTTGATGTCATTAATTAACCTTCATGATGTCGAAAAACAGTCTCTGTGACTCAACTC	2524	
Qy	2597	AAGCCTTGGTAAACCTTTTACAGCTCCAGCATCACTTATGACCTGCCCTCCAGTTCGGTGC	2656	
Db	2525	AAGCCTTGGTAAACCTTTTACAGCTCCAGCATCACTTATGACCTGCCCTCCAGTTCGGTGC	2584	
Qy	2657	CATCACCAGCTTATCCAGTGTGACTGAGAGCTGCCATCTTCTCTGTGGCAATTTGTCCCAT	2716	
Db	2585	CATCACCAGCTTATCCAGTGTGACTGAGAGCTGCCATCTTCTCTGTGGCAATTTGTCCCAT	2644	
Qy	2717	GTCT	2720	
Db	2645	GTCT	2648	

Search completed: July 13, 2005, 03:40:20
Job time : 1625 secs

This Page Blank (uspto)

1	Met	Asp	Thr	Ser	Arg	Leu	Gly	Val	Leu	Ser	Leu	Pro	Val	Leu	Gln	Leu	Ala	Thr	20
61	GGG	GGC	AGC	TCT	CCC	AGG	TCT	GGT	GTG	TCT	GCT	GAG	GGC	TCC	CCC	CAC	ACT	GTC	120
21	Gly	Gly	Ser	Pro	Arg	Ser	Gly	Val	Leu	Leu	Arg	Gly	Cys	Pro	Thr	His	Cys	His	40
121	GAG	CCC	CAC	GGC	AGC	AGT	TTG	TCT	CAG	GGT	TGAC	TGCT	CCG	ACC	TGG	GGC	TCT	CGG	180
41	Glu	Pro	Asp	Gly	Arg	Met	Leu	Leu	Arg	Val	Asp	Cys	Ser	Asp	Leu	Gly	Leu	Ser	60
181	CTT	TCC	AAC	CTC	AGC	GGT	TTT	CAC	CTC	CTT	ACG	CTC	TAG	ACCT	CAG	TAT	TGA	ACA	240
61	Pro	Ser	Asn	Leu	Ser	Val	Phe	Thr	Ser	Tyr	Leu	Asp	Leu	Ser	Met	Asn	Ile	Ser	80
241	CTG	TCT	CCG	GAAT	TCC	CTT	CCC	CCG	AGT	CCG	CTT	CTC	GCT	GAG	GAG	TAT	TAC	CTT	300
81	Leu	Leu	Pro	Asn	Pro	Leu	Pro	Ser	Leu	Arg	Phe	Leu	Glu	Leu	Arg	Leu	Ala	Gly	100
301	GCT	CTG	CAT	CAT	ATC	CCA	GGG	AGC	ATT	CAC	TGG	CCCT	TTT	CAC	AGT	CTT	TAA	GTT	360
101	Ala	Leu	Thr	Tyr	Ile	Pro	Lys	Gly	Ala	Phe	Thr	Gly	Leu	Tyr	Ser	Leu	Lys	Val	120
361	CTG	CAG	AAT	AAT	CAG	CTA	AGC	ACG	TAC	CCC	CAC	AGA	AGC	TCT	CGC	AAG	ATT	TGC	420
121	Leu	Gln	Asn	Asn	Gln	Leu	Arg	His	Val	Pro	Thr	Glu	Ala	Leu	Gln	Asn	Leu	Arg	140
421	CAAT	CCCT	CGG	TCT	GGA	TGCT	TAA	CCAC	CAT	CAG	CTAT	GTG	CCCC	CAAG	CTG	TTT	CAG	TGG	480
141	Gln	Ser	Leu	Arg	Leu	Asp	Ala	Asn	His	Ile	Ser	Tyr	Val	Pro	Pro	Ser	Cys	Phe	160
481	CTG	CAT	TCC	CTC	AGG	CAC	TGT	GCT	GGT	ATG	CAC	AAAT	CGT	TAA	CAG	AAA	TCC	CCG	540
161	Leu	His	Ser	Leu	Arg	His	Leu	Tyr	Phe	Asp	Asn	Ala	Leu	Thr	Glu	Ile	Pro	Val	180
541	GCT	TTT	TAGA	AGT	TTT	ATC	CGG	CAT	TGC	CAAG	CCAT	GAT	CAC	TTC	GGC	CTT	TGA	ACAAA	600
181	Ala	Phe	Arg	Ser	Leu	Ser	Ala	Leu	Gln	Ala	Met	Thr	Leu	Ala	Leu	Asn	Lys	Ile	200
601	ATAC	CAG	ACT	TAT	GCT	TTT	TGG	AACT	CTC	CAG	TTC	GTG	TAG	TCT	TAC	ATC	TCC	ACA	660
201	Ile	Pro	Asp	Tyr	Ala	Phe	Gly	Asn	Leu	Ser	Ser	Leu	Val	Val	Leu	His	Leu	His	220
661	AGA	TCC	CAC	TCC	CTG	GGA	AGAA	TAT	GCT	TTC	ATG	GGG	CTC	CAC	AGC	TAG	AGAC	TTC	720
221	Arg	Ile	His	Ser	Leu	Gly	Lys	Cys	Phe	Asp	Gly	Leu	His	Ser	Leu	Glu	Thr	Leu	240
721	TTA	AAAT	TACA	ATA	TAC	ATTT	CTT	ATG	ACA	TCC	CAAT	TGCA	ATT	AGG	CAC	ACT	TCT	CCA	780
241	Leu	Asn	Tyr	Asn	Asn	Leu	Asp	Glu	Phe	Pro	Thr	Ala	Ile	Arg	Thr	Leu	Ser	Asn	260
781	GAAC	TAG	ATT	TAT	CAC	ACA	CAAT	TAT	CAG	TCC	GNAT	CTC	TAG	AAAG	CAT	TTT	TAG	GCA	840
261	Glu	Leu	Gly	Phe	His	Ser	Asn	Asn	Ile	Arg	Ser	Ile	Pro	Glu	Lys	Ala	Phe	Val	280
841	CTT	TCT	CTT	TAT	TACA	ATA	TAC	ATTT	CTT	ATG	ACA	TCC	CAAT	TGCA	ATT	TGT	TGG	AGAC	900
281	Pro	Ser	Leu	Ile	Thr	Ile	His	Phe	Tyr	Asp	Asn	Pro	Ile	Gln	Phe	Val	Gly	Arg	300
901	TTT	CAAC	CAT	TTT	ACT	TGA	ACT	AAG	AC	ACT	GAT	CTC	TGA	AT	GGT	CGCT	CAC	AAA	960
301	Phe	Gln	His	Leu	Pro	Glu	Leu</												

Db	361	TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysLeuAsp	380
QY	1141	CTAAGACATAATGAAATCTACGAAATAAAGTTGACACTTTCACAGCAGTTGCTTAGCCCTC	1200
Db	381	LeuArgHisAsnGluLeuTyrGluLeuLysValAspThrPheGlnGlnLeuLeuSerLeu	400
QY	1201	CGATCGCTGAATTGGCTTGGAAACAAAAATGCTATTATTCACCCCAATGCATTTTCCACT	1260
Db	401	ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr	420
QY	1261	TTGGCATCCCTAATAAGACGTGGACCTATCGTCAACCTCTCTGTCGCTTTTCCATAACT	1320
Db	421	LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPhePheProIleThr	440
QY	1321	GGGTTACATCGTTAACTCACTTAAATTAACAGGAAATCATGCCTTACAGAGCTTGATA	1380
Db	441	GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle	460
QY	1381	TCATCTGAAAACCTTTCAGAACTCAAGGTTATAGAAATGCCTTTATGCTTACCAAGTGCTGT	1440
Db	461	SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCysCys	480
QY	1441	GCATTTGGAGTGTGTGAAATGCCTATAAGATTCTTAATCAATGGAATAAGGTGACAACT	1500
Db	481	AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn	500
QY	1501	AGCAGTATGACACACCTTCAATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560
Db	501	SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg	520
QY	1561	GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
Db	521	AspLeuGluAspPheLeuLeuAspPheGluGluAspLeuLysAlaLeuHisSerValGln	540
QY	1621	TGTTCACTTCCCAGGCCCTTCAAAACCCTGTGAACACCTGCTTGATGCTGCCTGCATC	1680
Db	541	CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle	560
QY	1681	AGAATTCGAGTGGGACCATAGCAGTCTCGCACTTACTTGTAATGCTTTGGTGACTTCA	1740
Db	561	ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer	580
QY	1741	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAAATGGGGTCATCGCA	1800
Db	581	ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla	600
QY	1801	GCAGTGAACATGCTCACGGAGTCTCCAGTCCGCTGCTGCTGCTGTGTGATGGTTCACT	1860
Db	601	AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr	620
QY	1861	TTTGGCAGCTTTCACGACATGTGCTGTGGGAGAAATGGGTGTGGTTCGCCATGTCAAT	1920
Db	621	PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle	640
QY	1921	GGTTTTTTTGCTTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980
Db	641	GlyPheLeuSerIlePheAlaSerGluSerSerValPheLeuLeuThrLeuAlaAlaLeu	660
QY	1981	GAGCGTGGGTTCTCTGTGAAATATTTCTGAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2040
Db	661	GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer	680
QY	2041	CTGAACTAATCATTTTGTCTGTGCCCTCTGCTGCTTACCACATGGCCGAGTTCCCTCG	2100
Db	681	LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaValProLeu	700
QY	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTTCCTTTTGGGAGCCC	2160
Db	701	LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro	720
QY	2161	AGCACCATGGGTACATGGTTCGTCAATCTTGCTCAATTCCTTTGCTTCCTCATGATG	2220
Db	721	SerThrMetGlyTrpMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet	740

Qy 2221 ACCATTGCTACCAAGCTCTACTGCAATTGGACAAAGGAGACCTGGAGATATTGG 2280
Db 741 ThrileAlaTyThrLysLeuTyCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Qy 2281 GACTGCTCTATGTTAAACACATTCGCTTGTCTTCCACCAACTGCATCTTAACCTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Qy 2341 CCTGTGGCTTCTTGTCTCTCTCTCTTATAAACCTTACATTTATCAGTCCCTGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Qy 2401 ATTAAGTTTATCTTCTGGTGTAGTCCCATCTCTCTGCATCTCAATCCCTCTCTAC 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTy 820
Qy 2461 ATCTGTTCATCTCTCACTTTAAGGAGATCTGGTGGCTGAGAAACCAACTACCTGTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuArgLysGlnThrTyVal 840
Qy 2521 TGGCAAGATCAAAACACCCCAAGCTTGATGTCAATTTAACTCTGTATGTGCAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Qy 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACCTTTTACCAGCTCCAGCATCACTTATGACCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerSerIleThrTyAspLeu 880
Qy 2641 CCTCCAGTTCGGTCCCATCACCAGCTTATCCAGTACGTAGAGCTGCATCTTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyProValThrGluSerCysHisLeuSerSer 900
Qy 2701 GTGGCATTGTGCCATGTCTC 2721
Db 901 ValAlaPheValProCysLeu 907
RESULT 2
JG0193
G protein-coupled receptor FEX - mouse
C:Species: Mus musculus (house mouse)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C:Accession: JG0193
R:Hermey, G.; Methner, A.; Schaller, H.C.; Hermans-Borgmeyer, I.
Biochem. Biophys. Res. Commun. 254, 273-279, 1999
A:Title: Identification of a novel seven-transmembrane receptor with homology to glycoprotein 135
A:Reference number: JG0193; MUID:99121227; PMID:9920770
A:Accession: JG0193
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-907 <HER>
Alignment Scores:
Pred. No.: 2,15e-43 Length: 907
Score: 51.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.62% Indels: 0
Gaps: 2
US-10-751-736-21 (1-2724) x JG0193 (1-907)
Qy 607 GACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAATAGATC 666
Db 203 AspTyAlaPheGlyAsnLeuSerSerLeuValValLeuHisLeuHisAsnArgIle 222
Qy 667 CACTCCCTGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGACATTTAGATTTAAAT 726
Db 223 HisSerLeuGlyLysCysPheAspGlyLeuHisSerLeuGluThrLeuAspLeuAsn 242
Qy 727 TACATAACCTTGATGAATTCCTCCACTGCAATT 759
Db 243 TyrAsnAsnLeuAspGluPheProThrAlaIle 253

RESULT 3
AF2638
conserved hypothetical protein Atu0506 [imported] - Agrobacterium tumefaciens (strain C58)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AF2638
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.B.; Chen, Y.; Woo, L.; Wood, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, S.E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AF2638
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-195 <KUR>
A:Cross-references: UNIPROT:Q8UHZ5; GB:AB008688; PIDN:AAL41524.1; PID:g17738854; GSPDB:G17738854
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0506
A:Map position: circular chromosome
Alignment Scores:
Pred. No.: 6.31 Length: 195
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
Gaps: 2
US-10-751-736-21 (1-2724) x AF2638 (1-195)
Qy 73 CCCAGGCTGCTGTGTTGCTGAGGGGC 99
Db 145 ProArgSerGlyValLeuLeuArgGly 153
RESULT 4
G97420
hypothetical protein AGR_C_894 [imported] - Agrobacterium tumefaciens (strain C58, Cerezo)
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: G97420
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: G97420
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-209 <KUR>
A:Cross-references: UNIPROT:Q8UHZ5; GB:AB007869; PIDN:AAK86320.1; PID:g15155438; GSPDB:G15155438
C:Genetics:
A:Gene: AGR C_894
A:Map position: circular chromosome
Alignment Scores:
Pred. No.: 6.25 Length: 209
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
Gaps: 2
US-10-751-736-21 (1-2724) x G97420 (1-209)
Qy 73 CCCAGGCTGCTGTGTTGCTGAGGGGC 99
Db 159 ProArgSerGlyValLeuLeuArgGly 167


```
RESULT 5
T22541
hypothetical protein F53B6.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T22541
R;White, S.
submitted to the EMBL Data Library, October 1996
A/Reference number: Z19578
A/Accession: T22541
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-342 <WIL>
A/Cross-references: UNIPROT:Q9XVW1; EMBL:Z81086; PIDN:CA803117.1; GSPDB:GNO0019; CESP:PS
A/Experimental source: clone F53B6
C/Genetics:
A/Gene: CESP:F53B6.4
A/Map position: 1
A/Introns: 70/3; 212/3; 314/3

Alignment Scores:
Pred. No.: 5.85 Length: 342
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.04% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x T22541 (1-342)
Qy 1592 TCCTCAAGTCAAGCAGCAATCTTCA 1566
Db 170 SerSerLysSerSerArgLysSerSer 178

RESULT 6
S06057
gene ND1 intron 3 protein 2 - Podospora anserina mitochondrion
C/Species: mitochondrion Podospora anserina
C/Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C/Accession: S06057
R;Cummings, D.J.; Domencico, J.M.; Michel, F.
Curr. Genet. 14, 253-264, 1988
A/Title: DNA sequence and organization of the mitochondrial ND1 gene from Podospora anse
A/Reference number: S06056; MUID:89063443; PMID:3197134
A/Accession: S06057
A/Molecule type: DNA
A/Residues: 1-580 <CUM>
A/Cross-references: UNIPROT:Q02714; EMBL:X13164
C/Genetics:
A/Genome: mitochondrion
A/Genetic code: SGC3
C/Keywords: mitochondrion

Alignment Scores:
Pred. No.: 5.46 Length: 580
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x S06057 (1-580)
Qy 1588 GAGGAAGACCTGAAGCCCTTCATCA 1614
Db 108 GluGluAspLeuLysAlaLeuHisSer 116

RESULT 7
A71400
probable disease resistance protein - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
A/Variety: Columbia
C/Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 09-Jul-2004
```

```
C/Accession: A71400
R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirks
P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giele
avanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A/Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech,
erhoft, A.; Moeres, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ansc
C.; Chalcwatzis, N.
A/Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thali
A/Reference number: A71400; MUID:98121113; PMID:9461215
A/Accession: A71400
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-869 <BEV>
A/Cross-references: UNIPROT:O32553; GB:Z97335; NID:G22444747; PID:e326857; PID:G2244748
C/Genetics:
A/Map position: 4COP9-4G3845

Alignment Scores:
Pred. No.: 5.17 Length: 869
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x A71400 (1-869)
Qy 1261 TTCCCATCCCTAATAAAGCTGACCTA 1287
Db 334 LeuProSerLeuIleLysLeuAspLeu 342

RESULT 8
S77257
hypothetical protein s11265 - Synechocystis sp. (strain PCC 6803)
C/Species: Synechocystis sp.
A/Variety: PCC 6803
C/Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 09-Jul-2004
C/Accession: S77257
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
o, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda,
DNA Res. 3, 109-136, 1996
A/Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.
A/Reference number: S74322; MUID:97061201; PMID:8905231
A/Accession: S77257
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-1978 <KAN>
A/Cross-references: UNIPROT:P73551; EMBL:D90907; GB:AB001339; NID:g1652618; PIDN:BAAL7591
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Alignment Scores:
Pred. No.: 4.63 Length: 1978
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.99% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x S77257 (1-1978)
Qy 997 CTGACTTAACTGGACACAGATCTCA 1023
Db 526 LeuThrLeuThrGlyAlaGlnIleSer 534

RESULT 9
S07957
hypothetical protein P-plus - Escherichia coli
C/Species: Escherichia coli
C/Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 08-Oct-1999
C/Accession: S07957
R;Plasterk, R.H.A.; van de Putte, P.
```


EMBO J. 4, 237-242, 1985
A:Title: The invertible P-DNA segment in the chromosome of *Escherichia coli*.
A:Reference number: S07185; MUID:85257443; PMID:3894006
A:Accession: S07957
A>Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-145 <PLA>
A:Cross-references: EMBL:X01805; NID:942407; PIDN:CAA25947.1; PID:G581189
C:Genetics:
A:Map position: 25 min
A:Start codon: GTG

Alignment Scores:
Pred. No.: 70.8 Length: 145
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
Gaps: 0
DB:

US-10-751-736-21 (1-2724) x S07957 (1-145)
Qy 988 GGTTCAGTTCAGTTCAATCAG 965
Db 128 GlyLeuGlnPheGlnLeuAsnGln 135

RESULT 10
AB1265
hypothetical protein lmol522 [imported] - *Listeria monocytogenes* (strain EGD-e)
C:Species: *Listeria monocytogenes*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB1265
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AB1265
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <GLA>
A:Cross-references: UNIPROT:Q93RD9; GB:NC_003210; PIDN:CAC99600.1; PID:G16410951; GSPDB: A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmol522
C:Superfamily: conserved hypothetical protein HI0670

Alignment Scores:
Pred. No.: 70.5 Length: 150
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
Gaps: 0
DB:

US-10-751-736-21 (1-2724) x AB1265 (1-150)
Qy 2136 GCAGAGGGGAGGCGCCATCTT 2113
Db 67 AlaGluArgGlyAlaIleLeu 74

RESULT 11
AD1627
hypothetical protein homolog lin157 [imported] - *Listeria innocua* (strain Clip11262)
C:Species: *Listeria innocua*
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 09-Jul-2004
C:Accession: AD1627
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, J.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, P.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AD1627
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-150 <GLA>
A:Cross-references: UNIPROT:Q92BJ1; GB:AL592022; PIDN:CAC96788.1; PID:G16414044; GSPDB: A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin1557
C:Superfamily: conserved hypothetical protein HI0670

Alignment Scores:
Pred. No.: 70.5 Length: 150
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
Gaps: 0
DB:

US-10-751-736-21 (1-2724) x AD1627 (1-150)
Qy 2136 GCAGAGGGGAGGCGCCATCTT 2113
Db 67 AlaGluArgGlyAlaIleLeu 74

RESULT 12
T31777
hypothetical protein F38H12.5 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31777
R:Woesner, J.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of *C. elegans* cosmid F38H12.
A:Reference number: Z21084
A:Accession: T31777
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-158 <WOE>
A:Cross-references: UNIPROT:O16345; EMBL:AF016436; PIDN:AAC25897.1; GSPDB:GN00023; CESP: A:Experimental source: strain Bristol N2; clone F38H12
C:Genetics:
A:Gene: CESP:F38H12.5
A:Map position: 5
A:Introns: 27/1; 38/3; 107/3

Alignment Scores:
Pred. No.: 70 Length: 158
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
Gaps: 0
DB:

US-10-751-736-21 (1-2724) x T31777 (1-158)
Qy 1226 TTGTTCCAGGCCAAATTCAGCGAT 1203
Db 151 LeuPheGlnAlaLysPheSerAsp 158

RESULT 13
F97440
hypothetical protein AGR_C_1201 [imported] - *Agrobacterium tumefaciens* (strain C58, Cere
C:Species: *Agrobacterium tumefaciens*
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: F97440
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* strain P2
A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: F97440
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-162 <KUR>
A;Cross-references: UNIPROT:Q8U5I8; GB:AB007869; PIDN:AAK86479.1; PID:g15155627; GSPDB:G000000000
A;Gene: AGR_C_1201
A;Map position: circular chromosome

Alignment Scores:
Pred. No.: 69.8 Length: 162
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x F97440 (1-162)

Qy 2068 CTGCTGGCCTTGACCATGGCGCA 2091

Db 22 LeuLeuAlaLeuThrMetAlaAla 29

RESULT 14

Dps family protein [imported] - *Caulobacter crescentus*
A;Accession: A87605
C;Species: *Caulobacter crescentus*
C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
R;Niernan, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Koln, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: A87605

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-178 <STO>

A;Cross-references: UNIPROT:Q9A4G1; GB:AB005673; NID:g13424489; PIDN:AAK24837.1; GSPDB:G000000000

C;Genetics:

A;Gene: CC2873

Alignment Scores:

Pred. No.: 68.9 Length: 178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x A87605 (1-178)

Qy 613 GCCTTGGAACTCTCCAGCTTG 636

Db 104 AlaPheGlyAenLeuSerSerLeu 111

RESULT 15

S75434

hypothetical protein c05017 - *Sulfolobus solfataricus*

C;Species: *Sulfolobus solfataricus*

C;Date: 11-Oct-1997 #sequence_revision 24-Oct-1997 #text_change 09-Jul-2004

R;Sensen, C.W.; Klenk, H.P.; Singh, R.K.; Allard, G.; Chan, C.C.Y.; Liu, Q.Y.; Penny, S. Mol. Microbiol. 22, 175-191, 1996

A;Title: Organizational characteristics and information content of an archaeal genome: 1

A;Reference number: S73076; MUID:97055432; PMID:8899719

A;Accession: S75434

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-179 <SEN>

A;Cross-references: UNIPROT:P96006; EMBL:Y08257; NID:g1707772; PID:e283869; PID:g1707840
A;Experimental source: strain P2
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1996
C;Genetics:
A;Start codon: GTG

Alignment Scores:
Pred. No.: 68.8 Length: 179
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.88% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x S75434 (1-179)

Qy 2035 TCTAGCCTGAAAGTAATCATTTTG 2058

Db 14 SerSerLeuLysValIleIleLeu 21

Search completed: July 13, 2005, 03:58:32

Job time : 83.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 20:27:05 ; Search time 8588 Seconds
(without alignments)
12073.483 Million cell updates/sec

Title: US-10-751-736-21

Perfect score: 2724
Sequence: 1 atgacacctccggctgg.....cattgtcccatgtctctaa 2724

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 34239544 seqs, 19032134700 residues

Word size : 0

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_ges1:*

9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2296	84.3	2724	9 AY411732	Homo sapi
2	951	34.9	1060	1 AL541959	AL541959
3	706	25.9	800	4 BG196708	BG196708
4	669	24.6	1104	1 AL530798	AL530798
5	620	22.8	671	4 BW795259	K-EST0077
6	538	19.8	590	6 CD617354	CD617354
7	529	19.4	873	5 BX403219	BX403219
8	475	17.4	475	5 BX107244	BX107244
9	472	17.3	472	1 AA460529	AA460529
10	418	15.3	556	5 BQ323949	BQ323949
11	408	15.0	2593	9 AY411733	AY411733
12	369	13.5	374	1 A1367723	A1367723
13	354	13.0	358	1 A1699236	A1699236
14	345	12.7	494	1 AA424098	AA424098
15	341	12.5	539	1 A1697103	A1697103
16	275	10.1	539	5 BX499910	BX499910
17	252	9.3	276	6 CD617347	CD617347
18	252	9.3	276	6 CD617349	CD617349
19	252	9.3	276	6 CD617353	CD617353
20	251	9.2	276	6 CD617346	CD617346
21	251	9.2	276	6 CD617348	CD617348
22	246	9.0	271	6 CD617352	CD617352
23	245	9.0	268	6 CD617342	CD617342
24	235	8.6	250	6 CD617356	CD617356

c 25	222	8.1	222	6	CD617357	CD617357	55041447J
c 26	222	8.1	222	6	CD617358	CD617358	55041471H
c 27	219	8.0	273	6	CD617343	55013050J	CD617343
c 28	201	7.4	275	6	CD617351	55013056J	CD617351
c 29	200	7.3	275	6	CD617350	55013056H	CD617350
c 30	193	7.1	330	5	BX357656	BX357656	BX357656
c 31	176	6.5	224	6	CD617359	55041471J	CD617359
c 32	176	6.5	432	8	AQ076504	CIT-HSP-2	AQ076504
c 33	169	6.2	275	6	CD617344	55013051H	CD617344
c 34	146	5.4	492	6	CD617355	55041415J	CD617355
c 35	139	5.1	397	1	A1539105	tp78b01.x	A1539105
c 36	125	4.6	542	1	A1962439	wq52c12.x	A1962439
c 37	119	4.4	396	1	AL113316	DKFp761C	AL113316
c 38	112	4.1	940	7	CK449409	892976 MA	CK449409
c 39	96	3.5	558	2	BE929414	NRO-GN002	BE929414
c 40	94	3.5	215	7	CO933257	AGENCOURT	CO933257
c 41	94	3.5	280	6	CD617345	55013051J	CD617345
c 42	90	3.3	391	8	B53762	CIT-HSP-201	B53762
c 43	87	3.2	324	1	A1206192	qr2xd02.x	A1206192
c 44	87	3.2	669	9	CE133614	tigr-ges-	CE133614
c 45	83	3.0	634	7	CF177761	806271 MA	CF177761

ALIGNMENTS

RESULT 1	AY411732	2724 bp	DNA	linear	GSS 16-DEC-2003
LOCUS	AY411732				
DEFINITION	Homo sapiens GPR49 gene, VIRTUAL TRANSCRIPT, partial sequence,				
ACCESSION	AY411732				
VERSION	AY411732.1	GI:39767700			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	1 (bases 1 to 2724)				
	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,				
	Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,				
	Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,				
	Adams, M.D. and Cargill, M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous				
JOURNAL	gene trios				
PUBMED	Science 302 (5652), 1960-1963 (2003)				
REFERENCE	14671302				
AUTHORS	2 (bases 1 to 2724)				
	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,				
	Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,				
	Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,				
	Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,				
COMMENT	Rockville, MD 20850, USA				
	This sequence was made by sequencing genomic exons and ordering				
	them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..2724				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
gene	<1..>2724				
	/gene="GPR49"				
	/locus_tag="HCM4322"				
ORIGIN					
Query Match	84.3%;	Score 2296;	DB 9;	Length 2724;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2296;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	429	CGGCTCGAGTCAACACATCAGCTATGTGCCCCAGCTGTTTCAGTCGCTGCATTC	488		

QY 2649 TTCCGTGCGCATCACCAGCTTATCCAGTGAAGTGGAGCTGCCATCTTTCTCTGTGGCAATT 2708
 Db 2649 TTCCGTGCGCATCACCAGCTTATCCAGTGAAGTGGAGCTGCCATCTTTCTCTGTGGCAATT 2708
 QY 2709 TGTCCCATGCTCTCTAA 2724
 Db 2709 TGTCCCATGCTCTCTAA 2724

RESULT 2
 LOCUS AL541959 1060 bp mRNA linear EST 24-MAR-2004
 DEFINITION AL541959 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE007YC19
 5-PRIME, mRNA sequence.

ACCESSION AL541959
 VERSION AL541959.3 GI:45717535
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:30546637.
 Contact: Genoscope

Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
 was not normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.

This sequence belongs to sequence cluster 7021.f

For more information about this cluster, see

http://www.genoscope.cns.fr/cdna?e=CS0DE007AB10QPI&c=7021.f.

FEATURES

Location/Qualifiers
 1..1060
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DE007YC19"
 /tissue_type="PLACENTA"
 /clone_lib="Homo sapiens PLACENTA"
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
 with a NotI-oligo(dT) primer. Five prime end enriched,
 double-strand cDNA was digested with Not I and cloned into
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.
 Library was not normalized."

ORIGIN

Query Match 34.9%; Score 951; DB 1; Length 1060;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 951; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 683 AATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGATTAAATACAAATACCTTGATG 742
 Db 1 AATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGATTAAATACAAATACCTTGATG 60
 QY 743 AATTCGCCACTGCAATTAGGACACTCTCCAACTTTAAAGAACTAGGATTTTCATAGCAACA 802
 Db 61 AATTCGCCACTGCAATTAGGACACTCTCCAACTTTAAAGAACTAGGATTTTCATAGCAACA 120
 QY 803 ATATCAGGTGATACCTGAGAAGCAATTTAGGCAACCTTCTCTTATTACATACATT 862
 Db 121 ATATCAGGTGATACCTGAGAAGCAATTTAGGCAACCTTCTCTTATTACATACATT 180
 QY 863 TCTATGACATCCCAATCAATTTGTTGGAGATCTGCTTTTCAACATTTTACCTGAACATA 922
 Db 181 TCTATGACATCCCAATCAATTTGTTGGAGATCTGCTTTTCAACATTTTACCTGAACATA 240

QY 923 GAACACTGACTCTGAATGGTGCCTCACAAATAAAGTAATTTCTTGATTTAACTGGAACCTG 982
 Db 241 GAACACTGACTCTGAATGGTGCCTCACAAATAAAGTAATTTCTTGATTTAACTGGAACCTG 300
 QY 983 CAAACCTGGAGAGTCTGACTTTTAACTGGAGACAGATCTCATCTCTTCTCTCAACCGTCT 1042
 Db 301 CAAACCTGGAGAGTCTGACTTTTAACTGGAGACAGATCTCATCTCTTCTCTCAACCGTCT 360
 QY 1043 GCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCTTACAACCTATTAGAAGATTTAC 1102
 Db 361 GCAATCAGTTACCTTAATCTCCAAGTGTAGATCTGTCTTACAACCTATTAGAAGATTTAC 420
 QY 1103 CCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATAGCACTAAGACATTAATGAATCTAGG 1162
 Db 421 CCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATAGCACTAAGACATTAATGAATCTAGG 480
 QY 1163 AAATTTAAAGTTTGACACTTTTCCAGCAGTGTCTTAGCTCCGATCGCTGAATTTGGCTTGA 1222
 Db 481 AAATTTAAAGTTTGACACTTTTCCAGCAGTGTCTTAGCTCCGATCGCTGAATTTGGCTTGA 540
 QY 1223 ACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACTTTGCCATCCCTTAATAAGCTGG 1282
 Db 541 ACAAAATTTGCTATTATTACCCCAATGCAATTTTCCACTTTGCCATCCCTTAATAAGCTGG 600
 QY 1283 ACCTATGCTCCAACTCTCTGTCTTCTTCTTAATAGTGGGTATCATGTTAACTCACT 1342
 Db 601 ACCTATGCTCCAACTCTCTGTCTTCTTCTTAATAGTGGGTATCATGTTAACTCACT 660
 QY 1343 TAAATAATTAACAGGAAATCATGCTTTACAGAGCTTGATATCATCTGAAAACCTTTCCAGAAC 1402
 Db 661 TAAATAATTAACAGGAAATCATGCTTTACAGAGCTTGATATCATCTGAAAACCTTTCCAGAAC 720
 QY 1403 TCAAGGTTTATAGAAATGCTTTATGCTTACCAAGTGTGTGCAATTTGGAGTGTGTGAGAATG 1462
 Db 721 TCAAGGTTTATAGAAATGCTTTATGCTTACCAAGTGTGTGCAATTTGGAGTGTGTGAGAATG 780
 QY 1463 CCTATAGATTTCTTAATCAATGGAATTAAGGTGACACAGCAGTATGACACCTTCATA 1522
 Db 781 CCTATAGATTTCTTAATCAATGGAATTAAGGTGACACAGCAGTATGACACCTTCATA 840
 QY 1523 AGAAGATGCTGGAAATGTTTCCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTG 1582
 Db 841 AGAAGATGCTGGAAATGTTTCCAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTG 900
 QY 1583 ACTTTGAGGAGAGACCTGAAAGCCCTTCATTCAAGTGCAGTGTTCACCTTTCCC 1633
 Db 901 ACTTTGAGGAGAGACCTGAAAGCCCTTCATTCAAGTGCAGTGTTCACCTTTCCC 951

RESULT 3

LOCUS BG196708 800 bp mRNA linear EST 21-APR-2001
 DEFINITION RST15934 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG196708
 VERSION BG196708.1 GI:13718395
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
 Cain, S., Leventhal, C., Thornton, M., Ramachandran, R.,
 Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
 Mays, R., Smith, E., Veloso, N., Klika, A., Hess, J., Cothren, K., Lo, K.,
 Offenbacher, J., Danzig, J. and Ducar, M.
 Creation of genome-wide protein expression libraries using random
 activation of gene expression

TITLE

Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL

21227151

MEDLINE

11329013

PUBMED

Contact: Scott J. Cain

Athersys, Inc.

3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atherays.com
High quality sequence stop: 491.

FEATURES	SOURCE
1. High Accuracy: The model achieves a high accuracy rate, indicating its effectiveness in predicting the target variable.	Model Performance Metrics
2. Robustness: The model demonstrates robustness against various input variations and noise, ensuring reliable predictions.	Model Performance Metrics
3. Interpretability: The model's predictions are interpretable, allowing users to understand the underlying factors influencing the outcome.	Model Performance Metrics
4. Scalability: The model is scalable, capable of handling large volumes of data and complex tasks efficiently.	Model Performance Metrics
5. Real-time Processing: The model supports real-time data processing, enabling immediate insights and decision-making.	Model Performance Metrics
6. Integration with Existing Systems: The model seamlessly integrates with existing data systems and workflows, facilitating easy adoption.	Model Performance Metrics
7. Customizable Parameters: The model offers customizable parameters, allowing users to tailor the model to specific requirements.	Model Performance Metrics
8. Comprehensive Documentation: The model is accompanied by comprehensive documentation, providing detailed information on its usage and capabilities.	Model Performance Metrics
9. Regular Updates: The model is regularly updated with the latest data and algorithms, ensuring its performance remains state-of-the-art.	Model Performance Metrics
10. Support and Training: The model is supported by a dedicated team, offering training and assistance to users.	Model Performance Metrics

ORIGIN

Query Match	25.9%;	Score 706;	DB 4;	Length 800;
Best Local Similarity	100.0%;	Prod. No. 0;		
Matches	706;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1210	AAATTGGCTTGGACAAAAATGCTATTATTATTCACCCCAATGCATTTTCCACTTTGGCAATCC	1269	
Db	1	AAATTGGCTTGGACAAAAATGCTATTATTATTCACCCCAATGCATTTTCCACTTTGGCAATCC	60	
Qy	1270	CTAATAAGCTGGACCTTATCGTCCAACTCCTGTGCTCTTTTCCCTATAAAGCTGGGTACAT	1329	
Db	61	CTAATAAGCTGGACCTTATCGTCCAACTCCTGTGCTCTTTTCCCTATAAAGCTGGGTACAT	120	
Qy	1330	GGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGACTTGATATCATCTGAA	1389	
Db	121	GGTTTAACTCACTTAAATTAACAGGAAATCATGCCCTTACAGACTTGATATCATCTGAA	180	
Qy	1390	AACTTTCCGAACCTCAAGGTTATAGAATGCTTATGCTTACCAGTGCCTGTGCATTTGGGA	1449	
Db	181	AACTTTCCGAACCTCAAGGTTATAGAATGCTTATGCTTACCAGTGCCTGTGCATTTGGGA	240	
Qy	1450	GTGTGTGAGAAATGCCTATAAGATTTCTTAATCAATGGAATAAAGGTGACAAACAGCAGTATG	1509	
Db	241	GTGTGTGAGAAATGCCTATAAGATTTCTTAATCAATGGAATAAAGGTGACAAACAGCAGTATG	300	
Qy	1510	GACGACCTTCANTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGTGACCTTGAA	1569	
Db	301	GACGACCTTCANTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGTGACCTTGAA	360	
Qy	1570	GAATTCCTGCTTGACTTTGAGGAGACCTCGAAAGCCCTTCATTCAGTGCAGTGTTCACCT	1629	
Db	361	GAATTCCTGCTTGACTTTGAGGAGACCTCGAAAGCCCTTCATTCAGTGCAGTGTTCACCT	420	
Qy	1630	TCCCGAGGCCCTTCAAAACCTGTGGAACACTGCTTGATGGCTGGCTGATCAGAATTTGGA	1689	
Db	421	TCCCGAGGCCCTTCAAAACCTGTGGAACACTGCTTGATGGCTGGCTGATCAGAATTTGGA	480	
Qy	1690	GTGTGGACCATAGCAGTTCTGGCACTTACTTGTAATGCTTTGTGTGACTTCAACAGTTTTTC	1749	
Db	481	GTGTGGACCATAGCAGTTCTGGCACTTACTTGTAATGCTTTGTGTGACTTCAACAGTTTTTC	540	
Qy	1750	AGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATGGGTTCATCGCAGCAGTGAAC	1809	
Db	541	AGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATGGGTTCATCGCAGCAGTGAAC	600	
Qy	1810	ATGCTCACGGAGTCTCCAGTGCCGTGCTGGCTGGTGTGATGGTTTCACTTTTGGCAGC	1869	
Db	601	ATGCTCACGGAGTCTCCAGTGCCGTGCTGGCTGGTGTGATGGTTTCACTTTTGGCAGC	660	
Qy	1870	TTTGACAGACATGGTGCCTGGTGGGAAATGGGGTTGGTGGCCATG	1915	
Db	661	TTTGACAGACATGGTGCCTGGTGGGAAATGGGGTTGGTGGCCATG	706	

RESULT 4	
ALU530798	1104 bp mRNA linear EST 24-MAR-2004
LOCUS	ALU530798 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION	cDNA clone CSDD008YF22 5-PRIME, mRNA sequence.
ACCESSION	ALU530798
VERSION	ALU530798.3 GI:45705748
KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1. (bases 1 to 1104) Li,W.B., Gruber,C., Jesse,J. and Polayes,D. Full-length cDNA libraries and normalization Unpublished (2001)
REFERENCE	On Feb 13, 2001 this sequence version replaced gi:31068631.
AUTHORS	Contact: Genoscope
TITLE	Genoscope - Centre National de Sequencage
JOURNAL	Genoscope - Centre National de Sequencage 2 rue Gaston Creneau, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
COMMENT	1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7021.f For more information about this cluster, see http://www.genoscope.cns.fr/cdna?s=CSDDP008PC110P1kc=7021.f .

FEATURES source

ORIGIN

Query Match	24.6%; Score 669; DB 1; Length 1104;
Best Local Similarity	99.6%; Pred. No. 0;
Matches	939; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
Qy	999 GACTTTAACTGGAGCACAGATCTCATCTCTTCCTCAAAACCGTCTGCNAATCAGTTACCTAA 1058
Db	1 GACTTTAACTGGAGCACAGATCTCATCTCTTCCTCAAAACCGTCTGCNAATCAGTTACCTAA 60
Qy	1059 TCTCCAAGTCTAGATCTGTCCTTCAACAACCTATTAGAAAGATTATCCACAGTTTTTTCAGTCTG 1118
Db	61 TCTCCAAGTCTAGATCTGTCCTTCAACAACCTATTAGAAAGATTATCCCAAGTTTTTCAGTCTG 120
Qy	1119 CCAAAAGCTTTCAGAAAAATTGACCTTAAGACATAATGAAATCTAGGAAATTAAGTTCACAC 1178
Db	121 CCAAAAGCTTCAGAAATTTGACCTTAAGACATAATGAAATCTAGGAAATTAAGTTCACAC 180
Qy	1179 TTTCCAGCAGTTGCTTTAGCCTCCGATCGCTGAAATTTGGCTTTGGAAACAAAATTTGCTATTAT 1238
Db	181 TTTCCAGCAGTTGCTTTAGCCTCCGATCGCTGAAATTTGGCTTTGGAAACAAAATTTGCTATTAT 240
Qy	1239 TCACCCCAATGCAATTTTCGACTTTGGCATCCCTTAATAAAGCTGGACCTATCGTCCCAACCT 1298
Db	241 TCACCCCAATGCAATTTTCCAATTTGGCATCCCTTAATAAAGCTGGACCTATCGTCCCAACCT 300
Qy	1299 CCTGTGCTCTTTTCCCTATAACTGGTTFACATGGTTTAACTCACTTAAAAATTAACACAGAAA 1358
Db	301 CCTGTGCTCTTTTCCCTATAACTGGTTFACATGGTTTAACTCACTTAAAAATTAACACAGAAA 360
Qy	1359 TCATGCCCTTACAGAGCTTGGATATCATCTGTAAGAACTTTTCAGAACTCAAGCTTATAGAAT 1418


```
Db 361 TCATGCTTACAGAGCTTGATATCATCTGAAAACCTTCCAGAACTCAAGCTTATAGAAAT 420
Qy 1419 GCCT-TATGCTTACCAGTCTGTGCAATTTGGAGTGTGAGAAATGCCATATAAGATTCTTA 1477
Db 421 GCCTATATGCTTACCAGTCTGTGCAATTTGGAGTGTGAGAAATGCCATATAAGATTCTTA 480
Qy 1478 ATCAATGGAATAAGGTGACACAGCATGATGAGACCTTTCATAGAAGATGCTGGA 1537
Db 481 ATCAATGGAATAAGGTGACACAGCATGATGAGACCTTTCATAGAAGATGCTGGA 540
Qy 1538 TGTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCCTGCTTGCATTTTGAGGAAGACC 1597
Db 541 TGTTTCAGGCTCAAGATGAACGTGACCTTGAAGATTTCCTGCTTGCATTTTGAGGAAGACC 600
Qy 1598 TGAAGACCCCTTCATTCAGTGCAGTGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAAC 1657
Db 601 TGAAGACCCCTTCATTCAGTGCAGTGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAAC 660
Qy 1658 ACCTGCTTGATGGCTGGCTGATCAGAAATTTGGAGTGTGAGACCATAGCAGTTCTGGCACTTA 1717
Db 661 ACCTGCTTGATGGCTGGCTGATCAGAAATTTGGAGTGTGAGACCATAGCAGTTCTGGCACTTA 720
Qy 1718 CTGTGTAATGCTTTGTGTGACTCAACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCACTTA 1777
Db 721 CTGTGTAATGCTTTGTGTGACTCAACAGTTTTTCAGATCCCTCTGTACATTTTCCCCCACTTA 780
Qy 1778 AACTGTTAATTTGGGTCATTCGAGCAGTGAAATGCTCAAGGAGTCTCCAGTGCCTGTC 1837
Db 781 AACTGTTAATTTGGGTCATTCGAGCAGTGAAATGCTCAAGGAGTCTCCAGTGCCTGTC 840
Qy 1838 TGGCTGGTGTGATGCGTTCATTTTGGCAGCTTTTGGCAGCATGCTGCTGGTGGGAGA 1897
Db 841 TGGCTGGTGTGATGCGTTCATTTTGGCAGCTTTTGGCAGCATGCTGCTGGTGGGAGA 900
Qy 1898 ATGGGCTTGGTGGCTGATGCTATGCTTTTGTGCTTTCCTATTTTTC 1940
Db 901 ATGGGCTTGGTGGCTGATGCTATGCTTTTGTGCTTTCCTATTTTTC 943

RESULT 5
BM795259 671 bp mRNA linear EST 05-MAR-2002
LOCUS K-EST007025 S21SNUS20 Homo sapiens cDNA clone S21SNUS20-31-C01 5',
DEFINITION mRNA sequence.
ACCESSION BM795259
VERSION BM795259.1 GI:19143491
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 671)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,O.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 31 row: C column: 01
High quality sequence stop: 671.
Location/Qualifiers
1..671
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S21SNUS20-31-C01"
/sex="F"

FEATURES
source
```

```
/tissue_type="Stomach"
/cell_type="Floating aggregates"
/cell_line="SNU-520"
/lab_host="Top10F/"
/clone_lib="S21SNUS20"
/note="Organ: Stomach; Vector: pTZ189pl; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
bacterial alkaline phosphatase (BAP) and then decapped
with tabacco acid pyrophosphatase (TAP). The decapped
intact mRNA was ligated with DNA-RNA linker including EcoR
I site by treatment of T4 RNA ligase and the first strand
cDNA was synthesized from oligo dt-selected mRNA by
priming with dt-tailed vector. The dt-tailed vector was
adjusted to have about 60nt. The cDNA vector was
circularized with E. coli DNA ligase after digestion of
EcoRI which site is also included in vector. An RNA strand
converted to a DNA strand by Okayama-Berg method. The
obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation method.
The cDNA libraries constructed by this method are
full-length enriched cDNA library."
ORIGIN
Query Match 22.8%; Score 620; DB 4; Length 671;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1775 TTAAACTGTTAATTTGGGTCATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGC 1834
Db 1 TTAAACTGTTAATTTGGGTCATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGC 60
Qy 1835 TGCTGGCTGGTGTGGATGCGTTCATTTTGGCAGCTTTTGCACGACATGCTGCTGGTGG 1894
Db 61 TGCTGGCTGGTGTGGATGCGTTCATTTTGGCAGCTTTTGCACGACATGCTGCTGGTGG 120
Qy 1895 AGAATGGGCTGGTGTGGATGCGTTCATTTTGGCAGCTTTTGTCCATTTTCTTCAGATCATCTG 1954
Db 121 AGAATGGGCTGGTGTGGATGCGTTCATTTTGTCCATTTTGTCCATTTTCTTCAGATCATCTG 180
Qy 1955 TTTTCTGCTTACTCTGCGAGCCCTGGAGCGTGGGTTCCTGTGAAATATTTCTGCAAAAT 2014
Db 181 TTTTCTGCTTACTCTGCGAGCCCTGGAGCGTGGGTTCCTGTGAAATATTTCTGCAAAAT 240
Qy 2015 TTGAAACGAAAGTCCATTTTCTAGCTGAAAGTAAATCATTTTGTGCTGCTGCTGCTG 2074
Db 241 TTGAAACGAAAGTCCATTTTCTAGCTGAAAGTAAATCATTTTGTGCTGCTGCTGCTG 300
Qy 2075 CTTTGACCATGGCGCAGTTCCTCTGCTGGGTGGCAGCAAGTATGGCGCTCCCTCTCT 2134
Db 301 CTTTGACCATGGCGCAGTTCCTCTGCTGGGTGGCAGCAAGTATGGCGCTCCCTCTCT 360
Qy 2135 GCCTGCCCTTTGCCCTTTTGGGAGCCACGACCATGGGTACATGCTGCTCTCATCTTTC 2194
Db 361 GCCTGCCCTTTGCCCTTTTGGGAGCCACGACCATGGGTACATGCTGCTCTCATCTTTC 420
Qy 2195 TCAATTCCTTTTGTCTTCTCATGATGACCATTTGCTACACCAAGCTCTACTGCAATTTGG 2254
Db 421 TCAATTCCTTTTGTCTTCTCATGATGACCATTTGCTACACCAAGCTCTACTGCAATTTGG 480
Qy 2255 ACAAGGAGACCTGGAGAAATATTTGGGACTGCTCTATGGTAAACACATTTGCCCTTTC 2314
Db 481 ACAAGGAGACCTGGAGAAATATTTGGGACTGCTCTATGGTAAACACATTTGCCCTTTC 540
Qy 2315 TCTTCACCAACTGCATCTCTAACTGCCCTGTGGCTTTCTGTCTCTCTCTCTCTTAAATA 2374
Db 541 TCTTCACCAACTGCATCTCTAACTGCCCTGTGGCTTTCTGTCTCTCTCTCTCTTAAATA 600
Qy 2375 ACCTTACATTTATCAGTCTGAAGTAATTAAGTTTATCTCTCTGCTGGTAGTCCCACTTC 2434
Db 601 ACCTTACATTTATCAGTCTGAAGTAATTAAGTTTATCTCTCTGCTGGTAGTCCCACTTC 660
Qy 2435 CTGCAATGCTC 2445
|||||
```



```
Db      661 CTGCATGCTC 671

RESULT 6
CD617354
LOCUS      590 bp      mRNA      linear      EST 12-JAN-2004
DEFINITION 55041415H1 PLP Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD617354
VERSION     CD617354.1 GI:40265619
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 590)
AUTHORS     Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE       Circular rapid amplification of cDNA ends for high-throughput
            extension cloning of partial genes
JOURNAL     Genomics 84 (1), 205-210 (2004)
COMMENT     Contact: Fu GK
            Incyte Genomics, Inc.
            3160 Porter Dr., Palo Alto, CA 94304, USA
            Tel: 6508454102
            Email: gfu@incyte.com.
FEATURES   Location/Qualifiers
            source          1..590
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone_lib="FLP"
                        /note="Vector: pDrive Cloning Vector"
ORIGIN
Query Match      19.8%; Score 538; DB 6; Length 590;
Best Local Similarity 100.0%; Pred. No. 2.8e-284;
Matches 538; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 292 GCGGGAACGCTCGACATACATCCCAAGGGAGCATTCAGTGGCCCTTACAGTCTTAAA 351
Db 16 GCGGGAACGCTCGACATACATTCCTCCAGGGAGCATTCAGTGGCCCTTACAGTCTTAAA 75
QY 352 GTTCTTATGCTGCAGAAATATCAGCTAAGACACAGTACCACAGAGCTCTGCAGAAATTG 411
Db 76 GTTCTTATGCTGCAGAAATATCAGCTAAGACACAGTACCACAGAGCTCTGCAGAAATTG 135
QY 412 CGAAGCCTTCATCCCTGCGTCTGGATGCTAACACATCAGCTATGTGCCCCAGCTGT 471
Db 136 CGAAGCCTTCATCCCTGCGTCTGGATGCTAACACATCAGCTATGTGCCCCAGCTGT 195
QY 472 TTCAGTGGCCTGCATTCCTCGAGGCACCTGTGGCTGGATGACATGCGTTAAACAGAAATC 531
Db 196 TTCAGTGGCCTGCATTCCTCGAGGCACCTGTGGCTGGATGACATGCGTTAAACAGAAATC 255
QY 532 CCCGTCCAGGCTTTTAGAAGTTTATCGGCATFTGCAAGCCATGACCTTTGGCCCTGAAACAAA 591
Db 256 CCCGTCCAGGCTTTTAGAAGTTTATCGGCATFTGCAAGCCATGACCTTTGGCCCTGAAACAAA 315
QY 592 ATACACACATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTC 651
Db 316 ATACACACATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTC 375
QY 652 CATAACATAGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCCTAGAG 711
Db 376 CATAACATAGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCCTAGAG 435
QY 712 ACTTTAGATTTAAATTACAATAAATCTTTGATGAATTTCCCACTGCAATTAGACACTCTCC 771
Db 436 ACTTTAGATTTAAATTACAATAAATCTTTGATGAATTTCCCACTGCAATTAGACACTCTCC 495
QY 772 AACCTTAAGNACTAGGATTTTCATAGCAACATATCAGGTGCGATACCTTGAGAAAGCAT 829
Db 496 AACCTTAAGNACTAGGATTTTCATAGCAACATATCAGGTGCGATACCTTGAGAAAGCAT 553
```

```
RESULT 7
BX403219
LOCUS      873 bp      mRNA      linear      EST 29-APR-2004
DEFINITION BX403219 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
            clone CSODI029YG09 5-PRIME, mRNA sequence.
ACCESSION  BX403219
VERSION     BX403219.2 GI:46875014
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   1 (bases 1 to 873)
AUTHORS     Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     On May 13, 2003 this sequence version replaced gi:30619059.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
            1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
            end enriched, double-strand cDNA was digested with Not I and cloned
            into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
            was normalized. Library was constructed by Life Technologies, a
            division of Invitrogen. This sequence belongs to sequence cluster
            10101.f
            For more information about this cluster, see
            http://www.genoscope.cns.fr/cdna?s=CS2AAW3ZE04_AW92A03_1&c=10101.f.
FEATURES   Location/Qualifiers
            source          1..873
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="CSODI029YG09"
                        /tissue_type="PLACENTA COT 25-NORMALIZED"
                        /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                        /note="1st strand cDNA was primed with a NotI-oligo(dT)
                        primer. Five prime end enriched, double-strand cDNA was
                        digested with Not I and EcoR V sites of the pCMVSPORT 6
                        vector. Library was normalized."
ORIGIN
Query Match      19.4%; Score 529; DB 5; Length 873;
Best Local Similarity 100.0%; Pred. No. 2.7e-279;
Matches 529; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAACACCTCCCGGCTCGGTGTGCTCTCTGCTTGCCTTGTGCTGCTGCTGCGAGCTGGCACC 60
Db 138 ATGGAACACCTCCCGGCTCGGTGTGCTCTCTGCTTGCCTTGTGCTGCTGCTGCGAGCTGGCACC 197
QY 61 GGGGGCAGCTCTCCAGGTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCATTC 120
Db 198 GGGGGCAGCTCTCCAGGTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCATTC 257
QY 121 GAGCCGACGCGCAGGATGTTGCTCAGGTGACCTGCTCCGACCTGGGGCTCTCGAGCTG 180
Db 258 GAGCCGACGCGCAGGATGTTGCTCAGGTGACCTGCTCCGACCTGGGGCTCTCGAGCTG 317
QY 181 CCTTCAACCTCAGCGCTTTTCACTCTACCTAGACCTCAGTATGAACAACATCAGTCAG 240
Db 318 CCTTCAACCTCAGCGCTTTTCACTCTCTCACTCTCCTAGACCTCAGTATGAACAACATCAGTCAG 377
QY 241 CTGCTCCCGAATCCCTGCGCCAGTCTCCGCTTCTCTGAGGAGTACCTCTTGGGGGAAC 300
Db 378 CTGCTCCCGAATCCCTGCGCCAGTCTCCGCTTCTCTGAGGAGTACCTCTTGGGGGAAC 437
QY 301 GCTCTGACATACATTCCTCAAGGAGGAGTTCACCTGGCCTTTTACAGTCTTAAAGTTCCTATG 360
Db 438 GCTCTGACATACATTCCTCAAGGAGGAGTTCACCTGGCCTTTTACAGTCTTAAAGTTCCTATG 497
QY 361 CTGCAGAAATACAGCTTAAGACACGTACCCACAGAGCTCTGCAGAAATTTTCGGAAGCCTT 420
```


Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1 (bases 1 to 2593)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2593)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..2593
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..2593
/gene="GPR49"
/locus_tag="HMC4322"
ORIGIN
Query Match 15.0%; Score 408; DB 9; Length 2593;
Best Local Similarity 99.1%; Pred. No. 1.8e-212;
Matches 758; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 1960 CTGCTTACTTGGGAGCCCTGGAGCGTGGTCTCTGTGGAATATTCGCAAAATTCGAA 2019
Db 1829 CTGCTTACTTGGGAGCCCTGGAGCGTGGTCTCTGTGGAATATTCGCAAAATTCGAA 1888
Qy 2020 ACGAAGCTCCATTTTCTAGCTGAAAGTAATCATTTTGTCTGTGCGCTCTGCGCCCTG 2079
Db 1889 ACGAAGCTCCATTTTCTAGCTGAAAGTAATCATTTTGTCTGTGCGCTCTGCGCCCTG 1948
Qy 2080 ACCATGGCGGAGTTCCCTGTGGTGGGAGCAAGTATGGCGCTCCCTCTCTGCGCTG 2139
Db 1949 ACCATGGCGGAGTTCCCTGTGGTGGGAGCAAGTATGGCGCTCCCTCTCTGCGCTG 2008
Qy 2140 CCTTTGCCCTTTGGGAGCCGACCATGGCTACATGGCTCGCTCATCTTGTGCTCAAT 2199
Db 2009 CCTTTGCCCTTTGGGAGCCGACCATGGCTACATGGCTCGCTCATCTTGTGCTCAAT 2068
Qy 2200 TCCCTTTGCTTCTCATGATGACCATTTGCTTACACCAAGCTTCTACTGCAATTTGGACAAG 2259
Db 2069 TCCCTTTGCTTCTCATGATGACCATTTGCTTACACCAAGCTTCTACTGCAATTTGGACAAG 2128
Qy 2260 GGAGACCTGGAGAATATTGGGAGCTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTC 2319
Db 2129 GGAGACCTGGAGAATATTGGGAGCTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTC 2188
Qy 2320 ACCAATGCATCTTAAACGCGCTGGCTTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2379
Db 2189 ACCAATGCATCTTAAACGCGCTGGCTTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 2248
Qy 2380 ACAATTTATCAGTCCTGAAGTAATTAAGTTTATCTCTTCTGGTGGTGGTGGTGGTGGTGG 2439
Db 2249 ACAATTTATCAGTCCTGAAGTAATTAAGTTTATCTCTTCTGGTGGTGGTGGTGGTGGTGG 2308
Qy 2440 TGTCTCAATCCCCTTCTCTACATCTTGTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2499
Db 2309 TGTCTCAATCCCCTTCTCTACATCTTGTTCATCTCTCTCTCTCTCTCTCTCTCTCTCT 2368
Qy 2500 CTGAGAAAGCAACCTAGCTCTGGGACAGATCAAAACACCCAGCTTGTATGTCAATTAAC 2559
Db 2369 CTGAGAAAGCAACCTAGCTCTGGGACAGATCAAAACACCCAGCTTGTATGTCAATTAAC 2428

Qy 2560 TCTGATGATGTCGAAAAACAGTCTCTGTGACCTCAACTCAAGCTTGGTAACCTTTTACCAGC 2619
Db 2429 TCTGATGATGTCGAAAAACAGTCTCTGTGACCTCAACTCAAGCTTGGTAACCTTTTACCAGC 2488
Qy 2620 TCCAGCATCACTTATGACCTGCTCCAGTTCCGTCGCCATCACCAGCTTATCCAGTGACT 2679
Db 2489 TCCAGCATCACTTATGACCTGCTCCAGTTCCGTCGCCATCACCAGCTTATCCAGTGACT 2548
Qy 2680 GAGAGCTGCCATCTTTCTCTGTGCGCATTTGTCCCATGTCTCTAA 2724
Db 2549 GAGAGCTGCCATCTTTCTCTGTGCGCATTTGTCCCATGTCTCTAA 2593
RESULT 12
AI367723/c
LOCUS AI367723 374 bp mRNA linear EST 15-FEB-1999
DEFINITION qv75h01.x1 NCI CGAP Ut1 Homo sapiens cDNA clone IMAGE:1987441.3;
similar to TR:Q90674 LUTEINIZING HORMONE RECEPTOR ;, mRNA
sequence.
ACCESSION AI367723.1 GI:4137468
VERSION AI367723
KEYWORDS EST
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 374)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 1807 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 307.
FEATURES
Location/Qualifiers
source
1..374
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1987441"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_hosts="DH10B"
/clone_lib="NCI CGAP Ut1"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"
ORIGIN
Query Match 13.5%; Score 369; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 4.6e-191;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2021 CGAAGCTCCATTTTCTAGCCTGAAAGTAATCTTCTCTGTGCGCTTGA 2080
Db 374 CGAAGCTCCATTTTCTAGCCTGAAAGTAATCTTCTCTGTGCGCTTGA 315
Qy 2081 CCATGGCGCAGTTCCCTCTGCTGGTGCGACAGATGCGCTCCCTCTCTGCGCTGC 2140
Db 314 CCATGGCGCAGTTCCCTCTGCTGGTGCGACAGATGCGCTCCCTCTCTGCGCTGC 255


```

Qy 2141 CTTTGCCCTTTGGGAGCCAGCACCATGGCTACATGGTGGCTCTCATCTTGCTCAATT 2200
    |||||
Db 254 CTTTGCCCTTTGGGAGCCAGCACCATGGCTACATGGTGGCTCTCATCTTGCTCAATT 195
    |||||
Qy 2201 CCCTTTGCTTCTCATGATGACCAATTCCTACACCAAGCTCTACTGCAATTTGGACAAGG 2260
    |||||
Db 194 CCCTTTGCTTCTCATGATGACCAATTCCTACACCAAGCTCTACTGCAATTTGGACAAGG 135
    |||||
Qy 2261 GAGACCTGGAGAAATTTGGGAGCTCTATATGGTAAACACATTCGCTTGCTCTTCA 2320
    |||||
Db 134 GAGACCTGGAGAAATTTGGGAGCTCTATATGGTAAACACATTCGCTTGCTCTTCA 75
    |||||
Qy 2321 CCRACTGCATCCTAAACTGCCCTGGCTTTCTTGCTCTCTCTCTTTAATAAACCTTGA 2380
    |||||
Db 74 CCAACTGCATCCTAAACTGCCCTGGCTTTCTTGCTCTCTCTCTTTAATAAACCTTGA 15
    |||||
Qy 2381 CATTATCA 2389
    |||||
Db 14 CATTATCA 6
    |||||

RESULT 13
AI699236/c
LOCUS
DEFINITION tx62f10.x1 NCI CGAP Utl1 Homo sapiens cDNA clone IMAGE:2274187 3',
similar to TR:075473 075473 ORPHAN G PROTEIN-COUPLED RECEPTOR HG38.
; mRNA sequence.
ACCESSION AI699236
VERSION tx62f10.x1
KEYWORDS EST.
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 358)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 3883 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 329.
Location/Qualifiers
FEATURES
source
1. 358
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2274187"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/clone_lib="NCI CGAP Utl1"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.75 kb. Life Technologies catalog #:
11538-014"

ORIGIN
Query Match 13.0%; Score 354; DB 1; Length 358;
Best Local Similarity 100.0%; Pred. No. 8.6e-183;
Matches 354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2044 AAAGTAATCATTTTGTCTGTGGCCCTGTGGCCCTTGACCATGGCGCAGTTCCTCCCTGCTG 2103

```

```

Db 358 AAAGTAATCATTTTGTCTGTGGCCCTGTGGCCCTTGACCATGGCGCAGTTCCTCCCTGCTG 299
    |||||
Qy 2104 GTTGGGAGCAAGTATGGGGCTCCCTCTCTGCTGCTGCTTGGCTTTGGGGAGCCAGC 2163
    |||||
Db 298 GGTGGGAGCAAGTATGGGGCTCCCTCTCTGCTGCTGCTTGGCTTTGGGGAGCCAGC 239
    |||||
Qy 2164 ACCATGGGCTACATGGTGGCTCTCATCTGCTCAATTTCCCTTGTCTTCATGATGACC 2223
    |||||
Db 238 ACCATGGGCTACATGGTGGCTCTCATCTGCTCAATTTCCCTTGTCTTCATGATGACC 179
    |||||
Qy 2224 ATTGCTTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAATATTGGGAC 2283
    |||||
Db 178 ATTGCTTACACCAAGCTCTACTGCAATTTGGACAAGGAGACCTGGAGAATATTGGGAC 119
    |||||
Qy 2284 TGCTCTATGTATAAACACATTCGCTGCTCTTCAACCACTGCATCTCAAACTGCCCT 2343
    |||||
Db 118 TGCTCTATGTATAAACACATTCGCTGCTCTTCAACCACTGCATCTCAAACTGCCCT 59
    |||||
Qy 2344 GTGGCTTTCTTGCTCTCTCTCTCTTAATAAACCTTACATTTATCAGTCTCGAA 2397
    |||||
Db 58 GTGGCTTTCTTGCTCTCTCTCTCTTAATAAACCTTACATTTATCAGTCTCGAA 5
    |||||

RESULT 14
AA424098 494 bp mRNA linear EST 16-OCT-1997
LOCUS zv80c01.r1 Soares total_fetus_Nb2HF8_9w Homo sapiens cDNA clone
DEFINITION IMAGE:759936 5', similar to TR:G1256414 G1256414 OVARIAN
FOLLICLE-STIMULATING HORMONE RECEPTOR. ; mRNA sequence.
ACCESSION AA424098
VERSION AA424098.1 GI:2103059
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 494)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,
White, Y., Wylie, T., Waterston, R., and Wilson, R.
WashU-Merck EST Project 1997
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -28ml3 rev2 Et from Amersham
High quality sequence stop: 471.
Location/Qualifiers
FEATURES
source
1. 494
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:759936"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/clone_lib="Soares total fetus Nb2HF8_9w"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5',
TGTTACCAATCGAAGTGGAGCGCGCTTAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

```


This page blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: July 13, 2005, 00:42:39 ; Search time 243.5 Seconds
(without alignments)
11457.115 Million cell updates/sec

Title: US-10-751-736-21

Perfect score: 907
Sequence: 1 atgagacactccggctgg.....catttgcccatgtctctaa 2724

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Word size: 1

Total number of hits satisfying chosen parameters: 3224408

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO pool/US10751736/runat 12072005 084327 23917/app query.fasta 1.2887
-DB=UniProt_03 -QW=Fastan -SUFFIX=oli_rup -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFWT=pto
-NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10751736 @CGN 1 1 369 @runat 12072005 084327 23917 -NCPUS=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : UniProt_03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	907	100.0	907	1 LGR5 HUMAN	O75473 homo sapien
2	51	5.6	363	2 Q8C8A7	Q8C8A7 mus musculus
3	51	5.6	907	1 LGR5 MOUSE	Q921P4 mus musculus
4	27	3.0	147	2 Q7IM96	Q7IM96 sus scrofa
5	18	2.0	828	1 LGR6 HUMAN	Q9HXB8 homo sapien
6	18	2.0	915	2 Q6UY15	Q6UY15 homo sapien
7	18	2.0	923	2 Q86VU0	Q86VU0 homo sapien
8	18	2.0	928	2 Q8BYD7	Q8BYD7 homo sapien
9	15	1.7	878	2 Q8BXS9	Q8BXS9 mus musculus
10	13	1.4	927	2 Q8N537	Q8N537 homo sapien
11	13	1.4	951	1 LGR4 HUMAN	Q9BXB1 homo sapien
12	13	1.4	951	1 LGR4 RAT	Q9Z2H4 rattus norv
13	11	1.2	134	2 Q8OT31	Q8OT31 mus musculus
14	11	1.2	162	2 Q8QUB8	Q8QUB8 mus musculus
15	11	1.2	459	2 Q8R301	Q8R301 mus musculus
16	9	1.0	168	2 Q8OZR7	Q8OZR7 mus musculus

c

c	17	9	1.0	187	2	Q8BYF2	Q8BYF2 mus musculus
	18	9	1.0	195	2	Q8UH25	Q8UH25 agrobacteri
	19	9	1.0	209	2	Q7D1B7	Q7D1B7 agrobacteri
	20	9	1.0	213	2	Q6ANT4	Q6ANT4 desulfotale
c	21	9	1.0	218	2	Q821I3	Q821I3 streptomyc
c	22	9	1.0	232	1	RH06 HUMAN	RH06 HUMAN
c	23	9	1.0	232	2	Q8BLR7	Q8BLR7 mus musculu
c	24	9	1.0	232	2	Q66JC6	Q66JC6 xenopus tro
c	25	9	1.0	232	2	Q7X477	Q7X477 xenopus lae
c	26	9	1.0	232	2	Q9W761	Q9W761 xenopus lae
c	27	9	1.0	342	2	Q9XVNI	Q9XVNI caenorhabdi
	28	9	1.0	369	2	Q93177	Q93177 caenorhabdi
	29	9	1.0	379	2	Q9DG05	Q9DG05 gallus gall
	30	9	1.0	424	2	Q9DG06	Q9DG06 gallus gall
	31	9	1.0	488	2	Q6PHA3	Q6PHA3 mus musculu
	32	9	1.0	589	2	Q9FHL8	Q9FHL8 arabidopsis
	33	9	1.0	610	2	Q7GEY4	Q7GEY4 podospora a
	34	9	1.0	638	2	Q6Z1I6	Q6Z1I6 oryza sativ
	35	9	1.0	653	2	Q8LKV9	Q8LKV9 aegilops ta
	36	9	1.0	657	2	Q6ZHK7	Q6ZHK7 oryza sativ
c	37	9	1.0	665	2	Q9F5F6	Q9F5F6 agrobacteri
	38	9	1.0	693	1	FSHR CHICK	FSHR CHICK
	39	9	1.0	700	2	Q02714	Q02714 podospora a
	40	9	1.0	869	2	Q23253	Q23253 arabidopsis
	41	9	1.0	871	2	Q7XNV7	Q7XNV7 oryza sativ
	42	9	1.0	879	2	Q68CI3	Q68CI3 oncorhynch
	43	9	1.0	891	2	Q93YT3	Q93YT3 arabidopsis
	44	9	1.0	999	2	Q8K7X8	Q8K7X8 streptococc
	45	9	1.0	1008	2	Q8P1F7	Q8P1F7 streptococc

ALIGNMENTS

RESULT 1

LGR5_HUMAN
ID LGR5_HUMAN STANDARD; PRT; 907 AA.

AC O75473; Q9UP75;

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)

DE Leucine-rich repeat-containing G protein-coupled receptor 5 precursor

DE (Orphan G protein-coupled receptor HG38) (G protein-coupled receptor

DE 49).

GN Name=GPR49; Synonyms=GPR67, LGR5;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98308104; PubMed=9642114; DOI=10.1006/bbrc.1998.8774;

RA McDonald T., Wang R., Bailey W., Xie G., Chen F., Caskey C.T., Liu Q.;

RT "Identification and cloning of an orphan G protein-coupled receptor of

RT the glycoprotein hormone receptor subfamily.";

RL Biochem. Biophys. Res. Commun. 247:266-270(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RA MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;

RT Hsu S.Y., Liang S.G., Hsueh A.J.W.;

RT "Characterization of two LGR genes homologous to gonadotropin and

RT thyrotropin receptors with extracellular leucine-rich repeats and a G

RT protein-coupled, seven-transmembrane region.";

RL Mol. Endocrinol. 12:1830-1845(1998).

CC -1- FUNCTION: Orphan receptor. It may be an important receptor for

CC signals controlling growth and differentiation of specific

CC embryonic tissues (by similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- TISSUE SPECIFICITY: Expressed in skeletal muscle, placenta, spinal

CC cord, and various region of brain.

CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

CC -1- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.

CC -----

Db 281 ProSerLeuIleThrIleHisPheThyAspAsnProIleGlnPheValGlyArgSerAla 300
Qy 901 TTTCAACATTACCTGAACAACTAAGAACTAGCTCTGAATGGTGGCTCAAAATAACTGAA 960
Db 301 PheGlnHisLeuProGluLeuArgThrLeuThrLeuAsnGlyAlaSerGlnIleThrGlu 320
Qy 961 TTTCTCGATTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTTGGAGCACAGATC 1020
Db 321 PheProAspLeuThrGlyThrAlaAsnLeuGluSerLeuThrLeuThrGlyAlaGlnIle 340
Qy 1021 TCATCTCTCTCAAAACCGTCTGCAATCAGATTACCTAATCTCCAAGTCTAGATCTGTCT 1080
Db 341 SerSerLeuProGlnThrValCysAsnGlnLeuProAsnLeuGlnValLeuAspLeuSer 360
Qy 1081 TACAACCTATTAGAGATTACCCAGTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db 361 TyrAsnLeuLeuGluAspLeuProSerPheSerValCysGlnLysLeuGlnLysIleAsp 380
Qy 1141 CTAAGACATAATGAATCTAGAAATTAAGTTAGACATTTCCAGCAGTTCTTAGCCTC 1200
Db 381 LeuArgHisAsnGluIleTyrGluIleLysValAspThrPheGlnGlnLeuSerLeu 400
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCGTATTATTACCCCAATGCAATTTTCCACT 1260
Db 401 ArgSerLeuAsnLeuAlaTrpAsnLysIleAlaIleIleHisProAsnAlaPheSerThr 420
Qy 1261 TTGCCATCCCTAATAAGCTGAGCACTATCGTCCAACTCTCGTCTTTTCCATAACT 1320
Db 421 LeuProSerLeuIleLysLeuAspLeuSerSerAsnLeuLeuSerPheProIleThr 440
Qy 1321 GGGTTACATGTTTAACTCATTAAATTACAGGAATCATGCTTACAGAGCTTGATA 1380
Db 441 GlyLeuHisGlyLeuThrHisLeuLysLeuThrGlyAsnHisAlaLeuGlnSerLeuIle 460
Qy 1381 TCATCTGAAACTTTCCAGAACTCAAGGTATAGAAATTCCTTATGCTTACCAGTCTGT 1440
Db 461 SerSerGluAsnPheProGluLeuLysValIleGluMetProTyrAlaTyrGlnCys 480
Qy 1441 GCATTTGGAGTGTGTGAGAATGCCTATAAGATTTTCTAATCAATGGAATAAAGGTGACAA 1500
Db 481 AlaPheGlyValCysGluAsnAlaTyrLysIleSerAsnGlnTrpAsnLysGlyAspAsn 500
Qy 1501 AGCAGTATGAGCAGCTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAAGCT 1560
Db 501 SerSerMetAspAspLeuHisLysLysAspAlaGlyMetPheGlnAlaGlnAspGluArg 520
Qy 1561 GACCTTGAAGATTTCCTGCTTGAATTTGAGAGAGACCTGAAAGCCCTTCATTGAGTGCAG 1620
Db 521 AspLeuGluAspPheLeuLeuAspPheGluAspLeuLysAlaLeuHisSerValGln 540
Qy 1621 TGTTCACTTCCCAGGCCCTTCAAAACCTGTGAAACACCTGCTTGTGATGGCTGGCTGATC 1680
Db 541 CysSerProSerProGlyProPheLysProCysGluHisLeuLeuAspGlyTrpLeuIle 560
Qy 1681 AGAATTGAGTGTGGACCATGACAGTCTGGCACTTACTTGTAATGCTTTGGTGACTTCA 1740
Db 561 ArgIleGlyValTrpThrIleAlaValLeuAlaLeuThrCysAsnAlaLeuValThrSer 580
Qy 1741 ACAGTTTTTCAGTCCCTCTGACATTTCCCAATTAACCTGTAATTTGGGGTTCATCGCA 1800
Db 581 ThrValPheArgSerProLeuTyrIleSerProIleLysLeuLeuIleGlyValIleAla 600
Qy 1801 CGAGTGAACATGCTCACGGAGTCTCCAGTGGCTGCTGGTGGTGGTGGTGGTGGTGGTGGT 1860
Db 601 AlaValAsnMetLeuThrGlyValSerSerAlaValLeuAlaGlyValAspAlaPheThr 620
Qy 1861 TTGGCAGCTTGGCAGCATGGTGGCTGGTGGGAGAAATGGGTGGTGGTGGTGGTGGTGGT 1920
Db 621 PheGlySerPheAlaArgHisGlyAlaTrpTrpGluAsnGlyValGlyCysHisValIle 640
Qy 1921 GGTTTTTGTCCATTTTGTCTCAGAACTCATCTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 641 GlyPheLeuSerIlePheAlaSerGluSerValPheLeuLeuThrLeuAlaLeu 660

Qy 1981 GAGCGTGGTCTCTCTGTGAATAATATCTGCAGAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 661 GluArgGlyPheSerValLysTyrSerAlaLysPheGluThrLysAlaProPheSerSer 680
Qy 2041 CTGAAAGTAAATCATTTTGTCTGTGGCCCTGCTGCTGACCATGGCCGCGAGTTCCTCCCTG 2100
Db 681 LeuLysValIleIleLeuLeuCysAlaLeuLeuAlaLeuThrMetAlaAlaValProLeu 700
Qy 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTCCCTTTGGGGAGAGCC 2160
Db 701 LeuGlyGlySerLysTyrGlyAlaSerProLeuCysLeuProLeuProPheGlyGluPro 720
Qy 2161 AGCAACATGGGCTACATGGTCTCTCATCTGTGCTCAATTCCTTTGCTTCTCATGATG 2220
Db 721 SerThrMetGlyTyrMetValAlaLeuIleLeuLeuAsnSerLeuCysPheLeuMetMet 740
Qy 2221 ACCATTGCCCTACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGAAATTTGG 2280
Db 741 ThrIleAlaTyrThrLysLeuTyrCysAsnLeuAspLysGlyAspLeuGluAsnIleTrp 760
Qy 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGTCTTTCACCAACTGCATCTCTAACTGC 2340
Db 761 AspCysSerMetValLysHisIleAlaLeuLeuLeuPheThrAsnCysIleLeuAsnCys 780
Qy 2341 CCTGTGCTTTTCTTGTCTCTCTCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Db 781 ProValAlaPheLeuSerPheSerSerLeuIleAsnLeuThrPheIleSerProGluVal 800
Qy 2401 ATTAAGTTTATCTCTCTGTTAGTCTCCACTCTCTGCTGCTCTCAATCCCTCTCTCTAC 2460
Db 801 IleLysPheIleLeuLeuValValProLeuProAlaCysLeuAsnProLeuLeuTyr 820
Qy 2461 ATCTTGTTCATCTCTCACTTTAAGGAGATCTGTGAGCCTGAGAAAGCAAACTACCTC 2520
Db 821 IleLeuPheAsnProHisPheLysGluAspLeuValSerLeuAsgLysGlnThrTyrVal 840
Qy 2521 TGGCAAGATGCAAAACACCAAGCTTGTGATGTCATTAATCTCTGATGATGTGCAAAAACAG 2580
Db 841 TrpThrArgSerLysHisProSerLeuMetSerIleAsnSerAspAspValGluLysGln 860
Qy 2581 TCCTGTGACTCACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCATTTATGACTCTG 2640
Db 861 SerCysAspSerThrGlnAlaLeuValThrPheThrSerSerIleThrTyrAspLeu 880
Qy 2641 CCTCCAGTTCCGGTCCATCACAGCTTATCCAGTGTGAGAGCTGCATCTTTCTCTCT 2700
Db 881 ProProSerSerValProSerProAlaTyrProValThrGluSerCysHisLeuSerSer 900
Qy 2701 GTGGCATTGTGCCATGCTCTC 2721
Db 901 ValAlaPheValProCysLeu 907

RESULT 2

Q8C8A7 ID Q8C8A7 PRELIMINARY; PRT; 363 AA.
AC Q8C8A7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched
DE library, clone: C130018C02 product: G protein-coupled receptor 49, full
DE insert sequence.
GN Name=Gpr49;
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=92279253; Pubmed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;

"The three subfamilies of leucine-rich repeat-containing G protein-coupled receptors (LGR): identification of LGR6 and LGR7 and the signaling mechanism for LGR7.";
Mol. Endocrinol. 14:1257-1271(2000).

[2]
SEQUENCE OF 406-828 FROM N.A.
MEDLINE=22040266; PubMed=12044878; DOI=10.1016/S0014-5793(02)02775-8;
Takeda S., Kadowaki S., Haga T., Takaesu H., Mitaku S.;
"Identification of G protein-coupled receptor genes from the human genome sequence.";
FEBS Lett. 520:97-101(2002).

[3]
SEQUENCE OF 431-828 FROM N.A.
TISSUE=Mammary gland;
PubMed=14702039; DOI=10.1038/ng1285;
Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R., Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H., Sekine M., Oabayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S., Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S., Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kubano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O., Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M., Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S., Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O., Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H., Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B., Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y., Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T., Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ikemori Y., Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikemori Y., Okamoto S., Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M., Togoashi T., Oyama M., Hata H., Watanabe M., Komatsu T., Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K., Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R., Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
"Complete sequencing and characterization of 21,243 full-length human cDNAs.";
Nat. Genet. 36:40-45(2004).

-1- FUNCTION: Orphan receptor.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
-1- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AF190501; AAG17168.1; -
EMBL; AB083616; BAB9329.1; -
EMBL; AK027377; BAB55071.1; ALT_INIT.
HSSP; P25147; 1D0B.
Genew; HGNC:19719; LGR6.

InterPro; IPR000276; GPCR_Rhodpsn.
InterPro; IPR002131; Gprhm_receptor.
InterPro; IPR001611; LRR.
InterPro; IPR003591; LRR_typ.
Pfam; PF00560; LRR; 12.
PRINTS; PR00373; GLYCHORMONER.
PRINTS; PR00237; GPCR_RHODOPSN.
PRINTS; PR00019; LEURICHRPT.

DR PROSITE; PS00237; G PROTEIN RECEPTOR FL1; FALSE NEG.
DR PROSITE; PS00262; G PROTEIN RECEPTOR FL2; FALSE NEG.
KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat; Transmembrane.

KW Transmembrane.
FT DOMAIN 1 428 Extracellular (Potential).
FT TRANSMEM 429 449 1 (Potential).
FT DOMAIN 450 459 Cytoplasmic (Potential).
FT TRANSMEM 460 480 2 (Potential).
FT DOMAIN 481 505 Extracellular (Potential).
FT TRANSMEM 506 526 3 (Potential).
FT DOMAIN 527 548 Cytoplasmic (Potential).
FT TRANSMEM 549 569 4 (Potential).
FT DOMAIN 570 588 Extracellular (Potential).
FT TRANSMEM 589 609 5 (Potential).
FT DOMAIN 610 635 Cytoplasmic (Potential).
FT TRANSMEM 636 656 6 (Potential).
FT DOMAIN 657 671 Extracellular (Potential).
FT TRANSMEM 671 691 7 (Potential).
FT DOMAIN 692 804 Cytoplasmic (Potential).
FT DOMAIN 813 816 Poly-Gly.
FT REPEAT 22 45 LRR 1.
FT REPEAT 46 69 LRR 2.
FT REPEAT 71 93 LRR 3.
FT REPEAT 94 117 LRR 4.
FT REPEAT 118 140 LRR 5.
FT REPEAT 142 164 LRR 6.
FT REPEAT 188 211 LRR 7.
FT REPEAT 212 236 LRR 8.
FT REPEAT 238 257 LRR 9.
FT REPEAT 258 281 LRR 10.
FT REPEAT 283 305 LRR 11.
FT DISULFID 503 578 By similarity.
FT CARBOHYD 15 15 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 34 34 N-linked (GlcNAc...) (Potential).
FT CONFLICT 406 410 CSPTP -> MISPT (in Ref. 2).
FT CONFLICT 628 628 W -> R (in Ref. 3).
FT CONFLICT 824 828 FASHV -> LLHTY (in Ref. 1).
SQ SEQUENCE 828 AA; 1B5971445AA2DBB4 CRC64;

Alignment Scores:
Pred. No.: 1.3e-08 Length: 828
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.98% Indels: 0
DB: 1 Gaps: 0

US-10-751-736-21 (1-2724) x LGR6_HUMAN (1-828)

QY 853 ACATACATTTCTATGACATCCATCCATTTGTTGGAGATCTGCTTTTCAA 906
|||||

Db 146 ThrilehisPheryrAspAsnProleGlnPheValGlyArgSerAlaPheGln 163
|||||

RESULT 6

ID Q6UY15 PRELIMINARY; PRT; 915 AA.
AC Q6UY15; 2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DB Gonadotropin receptor.
GN ORFNames=UN06427;
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=22897296; PubMed=12975309; DOI=10.1101/gr.1293003;
RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
Chen J., Chow B., Chui C., Crowley C., Curriel B., Deuel B., Dowd P.,
Baton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,

RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
RA Seshagiri S., Simmons L., Singh J., Smith V., Stinson J., Vagts A.,
RA Vandlen R., Watanabe C., Wileand D., Woods K., Xie M.H., Yansura D.,
RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
RA Godowski P.;
RT "The secreted protein discovery initiative (SPDI), a large-scale
RT effort to identify novel human secreted and transmembrane proteins: a
RT bioinformatics assessment";
RL Genome Res. 13:2265-2270 (2003).
DR EMBL; AY358119; AAQ8486.1; -.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016500; F: protein-hormone receptor activity; IEA.
DR GO; GO:0004872; F: receptor activity; IEA.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gprhm_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003885; LRR_cyst.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00365; LRR_SD22; 5.
DR SMART; SM00369; LRR_TYP; 14.
DR Receptor.
SQ SEQUENCE 915 AA; 99265 MW; D57DD7A9DBB555F4 CRC64;

Alignment Scores:
Pred. No.: 1.29e-08 Length: 915
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.98% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q6UV15 (1-915)

Qy 853 ACAATACATTCTATGACAAATCCCATCAATTTGTGGAGATCTGCTTTTCAA 906
|||||
Db 233 ThrileHisPheTyraAsnProIleGlnPheValGlyArgSerAlaPheGln 250
|||||

RESULT 7
Q86VUO
ID Q86VUO PRELIMINARY; PRT; 923 AA.
AC Q86VUO;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LGR6 protein (Fragment).
GN Name=LGR6;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heleth F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Faney J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalski U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC047905; AAH47905.1; -.
DR HSPF; Q9BZR6; IPR1.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0016500; F: protein-hormone receptor activity; IEA.
DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gprhm_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.
FT NON TER 1
SQ SEQUENCE 923 AA; 100047 MW; 541D6746DAB06813 CRC64;

Alignment Scores:
Pred. No.: 1.29e-08 Length: 923
Score: 18.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.98% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q86VUO (1-923)

Qy 853 ACAATACATTCTATGACAAATCCCATCAATTTGTGGAGATCTGCTTTTCAA 906
|||||
Db 241 ThrileHisPheTyraAsnProIleGlnPheValGlyArgSerAlaPheGln 258
|||||

RESULT 8
Q9BYD7
ID Q9BYD7 PRELIMINARY; PRT; 928 AA.
AC Q9BYD7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE VTS20631 (Fragment).
GN Name=VTS20631;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Okaze H., Hayashi A., Kozuma S., Saito T.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB049405; BAB39854.1; -.
DR HSPF; B25147; IPR0.
DR GO; GO:0016021; C: integral to membrane; IEA.
DR GO; GO:0007186; P: G-protein coupled receptor activity; IEA.
DR InterPro; IPR00276; GPCR_Rhodopsn.
DR InterPro; IPR002131; Gprhm_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00560; LRR_1; 15.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PRINTS; PR00019; LEURICHRPT.
DR SMART; SM00369; LRR_TYP; 8.


```

FT  NON TER      1      1
SQ  SEQUENCE    928 AA; 100487 MW; 4C3364ADEA89C463 CRC64;

Alignment Scores:
Pred. No.:      1.29e-08      Length:      928
Score:          18.00         Matches:    18
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     1.98%       Indels: 0
DB:              2           Gaps: 0

US-10-751-736-21 (1-2724) x Q9BYD7 (1-928)

QY  853 ACATATCATTTCTATGACATCCATCCATTTCTGGAGATCGTTTCAA 906
Db  246 ThrileHisPheTyAspAsnProIleInPheValGlyArgSerAlaPheGln 263

RESULT 9
Q8BX59 PRELIMINARY; PRT; 878 AA.
AC  Q8BX59
DT  01-MAR-2003 (TrEMBLrel. 23, Created)
DT  01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  Mus musculus adult retina cDNA, RIKEN full-length enriched library,
DE  clone.A930009A08 product:G PROTEIN-COUPLED RECEPTOR LGR4 homolog
DE  (Fragment);
GN  Name=Gpr48;
OS  Mus musculus (Mouse);
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX  NCBI_TaxID=10090;
RN  [1]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RX  MEDLINE=21085660; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA  Carninci P., Hayashizaki Y.;
RA  Kono H., Hayashizaki Y.;
RT  "High-efficiency full-length cDNA cloning.";
RL  Meth. Enzymol. 303:19-44(1999).
RN  [2]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RX  MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA  RIKEN FANTOM Consortium;
RT  "Functional annotation of a full-length mouse cDNA collection.";
RL  Nature 409:685-690(2001).
RN  [3]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RA  The RIKEN Consortium;
RA  the RIKEN Genome Exploration Research Group Phase I & II Team;
RT  "Analysis of the mouse transcriptome based on functional annotation of
RT  60,770 full-length cDNAs.";
RL  Nature 420:563-573(2002).
RN  [4]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RX  MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA  Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA  Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT  "Normalization and subtraction of cap-trapper-selected cDNAs to
RT  prepare full-length cDNA libraries for rapid discovery of new genes.";
RL  Genome Res. 10:1617-1630(2000).
RN  [5]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RX  MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA  Shibata K., Itoh M., Aizawa K., Nagao S., Sasaki N., Carninci P.,
RA  Kono H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA  Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA  Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA  Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA  Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,

```

```

RA  Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT  "RIKEN integrated sequence analysis (RISA) system-384-format
RT  sequencing pipeline with 384 multicapillary sequencer.";
RL  Genome Res. 10:1757-1771(2000).
RN  [6]
SEQUENCE FROM N.A.
RC  STRAIN=C57BL/6J; TISSUE=Retina;
RA  Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA  Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA  Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA  Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA  Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA  Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA  Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA  Saito R., Saitho H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA  Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA  Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,
RA  Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL  Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AK044357; BAC31882.1; -.
DR  HSSP; P25147; ID0B.
DR  MG; MGI:1891468; Gpr48.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR  GO; GO:0004872; F:receptor activity; IEA.
DR  GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR  InterPro; IPR000276; GPCR_Rhodpsn.
DR  InterPro; IPR002131; Gprhnm_receptor.
DR  InterPro; IPR001611; LRR.
DR  InterPro; IPR003591; LRR typ.
DR  InterPro; IPR007087; Znf_C2H2.
DR  Pfam; PF00001; 7tm 1; 1.
DR  Pfam; PF00560; LRR_1; 14.
DR  PRINTS; PR00373; GLYCHORMONER.
DR  PRINTS; PR00237; GPCRHHODPSN.
DR  PRINTS; PR00019; LEURICHRPT.
DR  SMART; SM00369; LRR_TYP; 5.
DR  PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR  PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW  Receptor.
FT  NON TER      1      1
SQ  SEQUENCE    878 AA; 96877 MW; 68E64B5EDEA11B37 CRC64;

Alignment Scores:
Pred. No.:      1.66e-05      Length:      878
Score:          15.00         Matches:    15
Percent Similarity: 100.00%    Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     1.65%       Indels: 0
DB:              2           Gaps: 0

US-10-751-736-21 (1-2724) x Q8BX59 (1-878)

QY  706 CTAGAGACTTTAGATTAAATTAACAATACCTTGATCAATCCCC 750
Db  154 LeuGluThrLeuAspLeuAsnTyAsnAsnLeuAspGluPhePro 168

RESULT 10
Q8N537 PRELIMINARY; PRT; 927 AA.
AC  Q8N537
DT  01-OCT-2002 (TrEMBLrel. 22, Created)
DT  01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT  01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE  GPR48 protein.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
SEQUENCE FROM N.A.
RP  TISSUE=Brain;
RC  MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX

```


RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullihy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences"; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

[2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC033039; AAB33039.1; -;
 DR HSSP; P25147; 1D0B.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR007087; Znf_C2H2.
 DR Pfam; PF00001; 7cm_1; 1.
 DR Pfam; PF01462; LRRNT; 1.
 DR Pfam; PF00560; LRR_1; 14.
 DR PRINTS; PR00373; GLYCHORMONER.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 5.
 DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_F1_2; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN 1.
 SQ SEQUENCE 927 AA; 101676 MW; C7B2F0C40E584CE8 CRC64;

Alignment Scores:
 Pred. No.: 0.00195 Length: 927
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.43% Indels: 0
 DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q8N537 (1-927)
 QY 622 AACCTCTCCAGCTGGTAGTCTACATCTCCATCAACAT 660
 Db 175 AsnLeuSerLeuValValLeuHisLeuHisAsnAsn 187
 RESULT 11
 LGR4_HUMAN
 ID_LGR4_HUMAN STANDARD; PRT; 951 AA.
 AC Q9BXL1; Q9NYD1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor
 DE (G protein-coupled receptor 48).
 GN Name=GPR48; Synonyms=LGR4;
 OS Homo sapiens (human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=21294803; PubMed=11401528; DOI=10.1006/bbrc.2001.4625;
 RA Loh E.D., Broussard S.R., Kolakowski L.F. Jr.;
 RT "Molecular characterization of a novel glycoprotein hormone G-protein-coupled receptor.";
 RL Biochem. Biophys. Res. Commun. 282:757-764 (2001).
 CC -1- FUNCTION: Orphan receptor.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in multiple steroidogenic tissues: placenta, ovary, testis and adrenal. Expressed also in spinal cord, thyroid, stomach, trachea, heart, pancreas, kidney, prostate and spleen.
 CC -1- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
 CC -1- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; AF346711; AAK31153.1; -;
 DR EMBL; AF346709; AAK31153.1; JOINED.
 DR EMBL; AF346710; AAK31153.1; JOINED.
 DR EMBL; AF257182; AAF68989.1; -;
 DR HSSP; Q9BZR6; 1OZN.
 DR MIN; 606666; -;
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR002131; Gphrmn_receptor.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR000372; LRR_Nterm.
 DR InterPro; IPR003591; LRR_TYP.
 DR Pfam; PF00001; 7cm_1; 1.
 DR Pfam; PF00560; LRR_15.
 DR Pfam; PF01462; LRRNT; 1.
 DR PRINTS; PR00373; GLYCHORMONER.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00013; LRRNT; 1.
 DR SMART; SM00369; LRR_TYP; 4.
 DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; FALSE_NEG.
 DR PROSITE; PS0262; G_PROTEIN_RECEPTOR_F1_2; 1.
 KW G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
 KW Signal; Transmembrane.
 FT SIGNAL 1 24 Potential.
 FT CHAIN 25 951 Leucine-rich repeat-containing G protein-coupled receptor 4.
 FT DOMAIN 25 544 Extracellular (Potential).
 FT TRANSMEM 545 565 1 (Potential).
 FT DOMAIN 566 575 Cytoplasmic (Potential).
 FT TRANSMEM 576 596 2 (Potential).
 FT DOMAIN 597 620 Extracellular (Potential).
 FT TRANSMEM 621 641 3 (Potential).
 FT DOMAIN 642 661 Cytoplasmic (Potential).
 FT TRANSMEM 662 682 4 (Potential).
 FT DOMAIN 683 703 Extracellular (Potential).
 FT TRANSMEM 704 724 5 (Potential).
 FT DOMAIN 725 756 Cytoplasmic (Potential).
 FT TRANSMEM 757 777 6 (Potential).
 FT DOMAIN 778 783 Extracellular (Potential).
 FT TRANSMEM 784 804 7 (Potential).
 FT DOMAIN 805 951 Cytoplasmic (Potential).
 FT REPEAT 55 79 LRR 1.
 FT REPEAT 81 103 LRR 2.
 FT REPEAT 104 127 LRR 3.


```
FT REPEAT 128 151 LRR 4.
FT REPEAT 153 175 LRR 5.
FT REPEAT 176 199 LRR 6.
FT REPEAT 201 223 LRR 7.
FT REPEAT 224 247 LRR 8.
FT REPEAT 248 270 LRR 9.
FT REPEAT 272 294 LRR 10.
FT REPEAT 318 341 LRR 11.
FT REPEAT 342 366 LRR 12.
FT REPEAT 368 387 LRR 13.
FT REPEAT 388 411 LRR 14.
FT REPEAT 413 435 LRR 15.
FT DISULFID 618 693 By similarity.
FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 294 294 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 505 505 N-linked (GlcNAc...) (Potential).
FT CONFLICT 292 292 P -> S (in Ref. 1; AAF68989).
FT CONFLICT 433 433 L -> P (in Ref. 1; AAF68989).
FT CONFLICT 668 668 L -> S (in Ref. 1; AAF68989).
SQ SEQUENCE 951 AA; 104460 MW; 5E0C2DFCF22CA1BB CRC64;

Alignment Scores:
Pred. No.: 0.00195 Length: 951
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 1 Gaps: 0

US-10-751-736-21 (1-2724) x LGR4_HUMAN (1-951)
Qy 622 AACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Dy 199 AsnLeuSerLeuValValLeuHisLeuHisAsnAen 211

RESULT 12
LGR4_RAT STANDARD; PRT; 951 AA.
AC Q922H4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 4 precursor.
GN Name=Gpr48; Synonyms=Lgr4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=99065210; PubMed=9849958; DOI=10.1210/me.12.12.1830;
RA Hsu S.Y., Liang S.-G., Haueh A.J.W.;
RT "Characterization of two LGR genes homologous to gonadotropin and
RT thyrotropin receptors with extracellular leucine-rich repeats and a G
RT protein-coupled, seven-transmembrane region.";
RL Mol. Endocrinol. 12:1830-1845(1998).
CC -!- FUNCTION: Orphan receptor.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; AF061443; AAC77910.1; --
```

```
DR HSP: Q9BZR6; LOZN.
DR RGD; 628615; Gpr48.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002131; Gphrmn_receptor.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000372; LRR_Nterm.
DR InterPro; IPR003591; LRR_typ.
DR Pfam; PF00001; 7tm_1; 1.
DR Pfam; PF00560; LRR; 15.
DR Pfam; PF01462; LRRNT; 1.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCR_Rhodopsin.
DR PRINTS; PR00019; LEURICHREPT.
DR SMART; SM00013; LRRNT; 1.
DR SMART; SM00369; LRR_TYP; 5.
DR PROSITE; PS00237; G_PROTEIN_RECPT_F1_1; FALSE_NEG.
DR PROSITE; PS00262; G_PROTEIN_RECPT_F1_2; 1.
DR G-protein coupled receptor; Glycoprotein; Leucine-rich repeat; Repeat;
KW Signal; Transmembrane.
KW SIGNAL 1 24 Potential.
FT CHAIN 25 951 Leucine-rich repeat-containing G protein-
FT FT Extracellular (Potential).
FT DOMAIN 25 544 1 (Potential).
FT TRANSMEM 545 565 Cytoplasmic (Potential).
FT DOMAIN 566 575 2 (Potential).
FT TRANSMEM 576 596 Extracellular (Potential).
FT DOMAIN 597 619 3 (Potential).
FT TRANSMEM 620 640 Cytoplasmic (Potential).
FT DOMAIN 641 661 4 (Potential).
FT TRANSMEM 662 682 Extracellular (Potential).
FT DOMAIN 683 703 5 (Potential).
FT TRANSMEM 704 724 Cytoplasmic (Potential).
FT DOMAIN 725 756 6 (Potential).
FT TRANSMEM 757 777 Extracellular (Potential).
FT DOMAIN 778 783 7 (Potential).
FT TRANSMEM 784 804 Cytoplasmic (Potential).
FT DOMAIN 805 951 LRR 1.
FT REPEAT 55 79 LRR 2.
FT REPEAT 81 103 LRR 3.
FT REPEAT 104 127 LRR 4.
FT REPEAT 128 151 LRR 5.
FT REPEAT 153 175 LRR 6.
FT REPEAT 176 199 LRR 7.
FT REPEAT 200 223 LRR 8.
FT REPEAT 225 247 LRR 9.
FT REPEAT 248 270 LRR 10.
FT REPEAT 272 294 LRR 11.
FT REPEAT 318 341 LRR 12.
FT REPEAT 342 366 LRR 13.
FT REPEAT 368 387 LRR 14.
FT REPEAT 388 411 LRR 15.
FT REPEAT 413 435 By similarity.
FT DISULFID 618 693 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 68 68 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 188 188 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 199 199 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 294 294 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 314 314 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 505 505 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 951 AA; 104138 MW; EDD56AC072123461 CRC64;

Alignment Scores:
Pred. No.: 0.00195 Length: 951
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservat: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.43% Indels: 0
DB: 1 Gaps: 0

US-10-751-736-21 (1-2724) x LGR4_RAT (1-951)
Qy 622 AACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Dy 622 AACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
|||||
```



```

Db      199 AsnLeuSerLeuValValLeuHisLeuHisAsnAen 211
RESULT 13
Q80T31 ID Q80T31 PRELIMINARY; PRT; 134 AA.
AC Q80T31,
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE G protein-coupled receptor GPR48 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;
RA Vassilatis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
RA Bergmann J.E., Gaitanaris G.A.;
RT "The G protein-coupled receptor repertoires of human and mouse.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).
DR EMBL; AY255619; AA085131.1; -
DR HSP; P25146; 106T.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR003591; LRR_TYP.
DR Pfam; PF00560; LRR_1; 4.
DR PRINTS; PR00019; LEURICRPT.
DR SMART; SM00369; LRR_TYP; 2.
KW Receptor.
FT NON TER 1 1
FT NON TER 134 134
SQ SEQUENCE 134 AA; 14851 MW; 4D355900D69C67E CRC64;

Alignment Scores:
Pred. No.: 0.289 Length: 134
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q80T31 (1-134)

Qy      418 CTTCAATCCCGCTCGTCAATCCATCCATC 450
|||||
Db      69 LeuGlnSerLeuArgLeuAspAlaAsnHisIle 79
|||||

RESULT 14
Q80UB8 ID Q80UB8 PRELIMINARY; PRT; 162 AA.
AC Q80UB8,
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Leucine-rich repeat-containing G protein-coupled receptor 6
DE (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;
RA Vassilatis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
RA Bergmann J.E., Gaitanaris G.A.;
RT "The G protein-coupled receptor repertoires of human and mouse.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).
DR EMBL; AY255562; AA085074.1; -
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.

```

```

FT NON TER 1 1
FT NON TER 162 162
SQ SEQUENCE 162 AA; 17608 MW; C593128C551FB824 CRC64;

Alignment Scores:
Pred. No.: 0.283 Length: 162
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q80UB8 (1-162)

Qy      2428 CCACCTTCCTGCATGCTCAATCCCTCTCTCTAC 2460
|||||
Db      70 ProLeuProAlaCysLeuAenProLeuLeuTyr 80
|||||

RESULT 15
Q8R301 ID Q8R301 PRELIMINARY; PRT; 459 AA.
AC Q8R301,
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lgr6 protein (Fragment).
GN Name=Lgr6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Czech II; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marta M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Czech II; TISSUE=Mammary tumor;
RA Strausberg R.;
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026896; AAH26896.1; -.
DR MGD; MGI:2441805; Lgr6.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0016500; F:protein-hormone receptor activity; IEA.
DR GO; GO:0007186; F:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR002131; Gphrmn_receptor.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00373; GLYCHORMONER.
DR PRINTS; PR00237; GPCRHHODPSN.
FT NON TER 1 1
FT NON TER 459 AA; 47889 MW; F0100BF073B81762 CRC64;
SQ SEQUENCE 459 AA; 47889 MW; F0100BF073B81762 CRC64;

```


Alignment Scores:
Pred. No.: 0.251 Length: 459
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.21% Indels: 0
DB: 2 Gaps: 0

US-10-751-736-21 (1-2724) x Q8R301 (1-459)

Qy 2428 CCACTTCCTGCATGCTCAATCCCTTCTCTAC 2460
Db 309 ProLeuProLaCysLeuAsnProLeuLeuTyr 319

Search completed: July 13, 2005, 03:56:52
Job time : 284.5 secs

181	Db	CTTCCAACTCAGCGTCTTCACTCTCTACCTCTAGACCTCAGTAGAAACAACATCAGTCAG	240
241	QY	CTGCTCCGGAATCCCTCGGCCAGTCTCCGCTCTCTGGAGGAGTTACGCTTTGCGGGAAAC	300
241	Db	CTGCTCCGGAATCCCTCGGCCAGTCTCCGCTCTCTGGAGGAGTTACGCTTTGCGGGAAAC	300
301	QY	GCTCTGACATACATTTCCCAAGGAGCATTCACTGGCCCTTTACAGCTTTAAAGTTCTTTATG	360
301	Db	GCTCTGACATACATTTCCCAAGGAGCATTCACTGGCCCTTTACAGCTTTAAAGTTCTTTATG	360
361	QY	CTGCAGAAATACAGTCTAAGACAAGTACCAAGCTCTGCAGAAATTTGCGAAGCCTT	420
361	Db	CTGCAGAAATACAGTCTAAGACAAGTACCAAGCTCTGCAGAAATTTGCGAAGCCTT	420
421	QY	CAATCCCTCGCTCTGGATGCTAAACAATCAGCTATGTGCCCCCAAGCTTTTCAGTGGC	480
421	Db	CAATCCCTCGCTCTGGATGCTAAACAATCAGCTATGTGCCCCCAAGCTTTTCAGTGGC	480
481	QY	CTGCATTTCCCTGAGGCACCTCTGGCTGGATGACAATCGTTTAAACAGAAATCCCGTCCAG	540
481	Db	CTGCATTTCCCTGAGGCACCTCTGGCTGGATGACAATCGTTTAAACAGAAATCCCGTCCAG	540
541	QY	GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTTGAAACAAATACACAC	600
541	Db	GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTTGAAACAAATACACAC	600
601	QY	ATACGAGACTATGCTTTGGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAAACAAT	660
601	Db	ATACGAGACTATGCTTTGGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAAACAAT	660
661	QY	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT	720
661	Db	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTAGAGACTTTAGAT	720
721	QY	TTAAATTAACAATAACCTTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAAA	780
721	Db	TTAAATTAACAATAACCTTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACCTTAAA	780
781	QY	GAACTAGGATTTATAGCAACAATATCAGGTCGATACCTGAGAAGCAATTTGTAGGCAAC	840
781	Db	GAACTAGGATTTATAGCAACAATATCAGGTCGATACCTGAGAAGCAATTTGTAGGCAAC	840
841	QY	CTTCTCTTATTACAAATACATTTCTATGACAATCCCATCCAAATTTGTTGGAGACTCTGCT	900
841	Db	CTTCTCTTATTACAAATACATTTCTATGACAATCCCATCCAAATTTGTTGGAGACTCTGCT	900
901	QY	TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCTCACAAATAACTGAA	960
901	Db	TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCTCACAAATAACTGAA	960
961	QY	TTTCTCTGATTTAACTGGAACCTGGAACCTGAGAGTCTGACTTTAACTGGAGCAAGATC	1020
961	Db	TTTCTCTGATTTAACTGGAACCTGGAACCTGAGAGTCTGACTTTAACTGGAGCAAGATC	1020
1021	QY	TCATCTCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGCT	1080
1021	Db	TCATCTCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGCT	1080
1081	QY	TACAACCTTATTAGAAGATTTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC	1140
1081	Db	TACAACCTTATTAGAAGATTTTACCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC	1140
1141	QY	CTAAGACATAATGAAATCTAGAAATTTAAAGTTGACACTTCCAGCAAGTTGCTTAGGCTC	1200
1141	Db	CTAAGACATAATGAAATCTAGAAATTTAAAGTTGACACTTCCAGCAAGTTGCTTAGGCTC	1200
1201	QY	CGATCGCTGAATTTGGCTTTGAAACAAATTTGCTATTTATCCACCAATGCAATTTCCACT	1260
1201	Db	CGATCGCTGAATTTGGCTTTGAAACAAATTTGCTATTTATCCACCAATGCAATTTCCACT	1260
1261	QY	TTGGCATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTCTGTCGCTCTTTTCTTAACT	1320

13261	TTGCGATCCCTAATAAAGCTGGAGCCTATACGTCCAACTCCTCTGTCGTCTTTTCTCTATAACT	13220
QY	GGGTTACATCGTGTAACTCACTTAAATAAAGAGAAATCATGCCCTTACAGAGCTTTGATA	1380
DB	GGGTTACATCGTGTAACTCACTTAAATTAACAGGAAATCATGCGCTTACAGAGCTTTGATA	1380
QY	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAAATGCGCTTATGCCATCAAGTGCCTGT	1440
DB	TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAAATGCGCTTATGCCATCAAGTGCCTGT	1440
QY	GCATTTGGAGTGTGAGAAATGCGCTATAGATTTCTAATCAATGGAAATTAAGGTGACAAC	1500
DB	GCATTTGGAGTGTGAGAAATGCGCTATAGATTTCTAATCAATGGAAATTAAGGTGACAAC	1500
QY	AGCAGTATGACGACCTTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560
DB	AGCAGTATGACGACCTTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560
QY	GACCTTGAAGATTTCTCTGTGACTTTTGAGGAAAGCCTGAAAGCCCTTCATTCAGTGCAG	1620
DB	GACCTTGAAGATTTCTCTGTGACTTTTGAGGAAAGCCTGAAAGCCCTTCATTCAGTGCAG	1620
QY	TGTTCCACTTCCCAGGCCCTTCAAAACCTGTGAACACCTGCTGATGGCTGGCTGATC	1680
DB	TGTTCCACTTCCCAGGCCCTTCAAAACCTGTGAACACCTGCTGATGGCTGGCTGATC	1680
QY	AGAATTGGAGTGTGGACCATAGCAGTCTTGCGCACTTACTTGTAATGCTTTGGTGACTTCA	1740
DB	AGAATTGGAGTGTGGACCATAGCAGTCTTGCGCACTTACTTGTAATGCTTTGGTGACTTCA	1740
QY	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATGGGGTCATCGCA	1800
DB	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCATTTAAACTGTTAATGGGGTCATCGCA	1800
QY	GCAGTGAACATGCTCAGCGGAGTCTCAGTGCCTGCTGCTGCTGCTGATGCGTTCACT	1860
DB	GCAGTGAACATGCTCAGCGGAGTCTCAGTGCCTGCTGCTGCTGCTGATGCGTTCACT	1860
QY	TTTGGCAGCTTTCACAGACATGCTGCTGCTGGGAGAAATGGGGTGTGGTGCATGTCATT	1920
DB	TTTGGCAGCTTTCACAGACATGCTGCTGCTGGGAGAAATGGGGTGTGGTGCATGTCATT	1920
QY	GGTTTTTTPGCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980
DB	GGTTTTTTPGCCATTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980
QY	GAGCGTGGGTTCTCTGTGAAATATTTCTGMAAATTTTGAAACGAAAGCTCCATTTTCTAGC	2040
DB	GAGCGTGGGTTCTCTGTGAAATATTTCTGMAAATTTTGAAACGAAAGCTCCATTTTCTAGC	2040
QY	CTGAAAGTAATCAATTTTGTCTCTGCGCTGCTGCGCTGTGACCAATGGCCGAGTTC	2100
DB	CTGAAAGTAATCAATTTTGTCTCTGCGCTGCTGCGCTGTGACCAATGGCCGAGTTC	2100
QY	CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGGGGAGCCCC	2160
DB	CTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGGGGAGCCCC	2160
QY	AGCACCATGGGCTACATGCTGCTCATCTTCTGCTCAATTCCTTTGCTTCTCATGATG	2220
DB	AGCACCATGGGCTACATGCTGCTCATCTTCTGCTCAATTCCTTTGCTTCTCATGATG	2220
QY	ACCATTCGCCCTACACCAAGCTCTACTGCAATTTTGGACAAGGAGACCTGGAGAAATTTGG	2280
DB	ACCATTCGCCCTACACCAAGCTCTACTGCAATTTTGGACAAGGAGACCTGGAGAAATTTGG	2280
QY	GACTGCTCTATGTTAAACACATATGGCCCTGTTGCTCTTCCACCAATCGCATCTCTAAACTGC	2340
DB	GACTGCTCTATGTTAAACACATATGGCCCTGTTGCTCTTCCACCAATCGCATCTCTAAACTGC	2340
QY	CTGTGCGCTTCTGTGCTTCTCTCTTTTAATAAACCTTTACATTTATCATGCTCTGAAGTA	2400
DB	CTGTGCGCTTCTGTGCTTCTCTCTTTAATAAACCTTTACATTTATCATGCTCTGAAGTA	2400

1321	Db		GGGTTACATGGTTTTAACTCACTTAAAAATAACAGGAATCATGCCTTACAGAGCTTGATA	1380
1381	Qy	TCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGCCTG	1440	TCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGCCTG
1381	Db	TCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGCCTG	1440	TCATCTGAAACATTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGCCTG
1441	Qy	GCATTTGGAGTGTGAGAAATGCTTATAAGATTTCTAATCAATGGAATAAAGGTGCAAC	1500	GCATTTGGAGTGTGAGAAATGCTTATAAGATTTCTAATCAATGGAATAAAGGTGCAAC
1441	Db	GCATTTGGAGTGTGAGAAATGCTTATAAGATTTCTAATCAATGGAATAAAGGTGCAAC	1500	GCATTTGGAGTGTGAGAAATGCTTATAAGATTTCTAATCAATGGAATAAAGGTGCAAC
1501	Qy	AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560	AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT
1501	Db	AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560	AGCAGTATGGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT
1561	Qy	GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGGAAGCCCTCATTCAGTGCAG	1620	GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGGAAGCCCTCATTCAGTGCAG
1561	Db	GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGGAAGCCCTCATTCAGTGCAG	1620	GACCTTGAAGATTTCTGCTGTGACTTTGAGGAAGACCTGGAAGCCCTCATTCAGTGCAG
1621	Qy	TGTTTCACTTCCCGAGGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTGGCTGATC	1680	TGTTTCACTTCCCGAGGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTGGCTGATC
1621	Db	TGTTTCACTTCCCGAGGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTGGCTGATC	1680	TGTTTCACTTCCCGAGGCCCTTCAAACCCCTGTGAACACCTGCTTGATGCTGGCTGATC
1681	Qy	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA
1681	Db	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740	AGAATTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA
1741	Qy	ACAGTTTTTCAGATCCCTCTGTATCACTTTCCCCCATTAACATGTTAATGGGTGATCGCA	1800	ACAGTTTTTCAGATCCCTCTGTATCACTTTCCCCCATTAACATGTTAATGGGTGATCGCA
1741	Db	ACAGTTTTTCAGATCCCTCTGTATCACTTTCCCCCATTAACATGTTAATGGGTGATCGCA	1800	ACAGTTTTTCAGATCCCTCTGTATCACTTTCCCCCATTAACATGTTAATGGGTGATCGCA
1801	Qy	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGTGCTGGTGTGATGGTTCACCT	1860	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGTGCTGGTGTGATGGTTCACCT
1801	Db	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGTGCTGGTGTGATGGTTCACCT	1860	GCAGTGAACATGCTCACGGAGTCTCCAGTGCCTGTGCTGGTGTGATGGTTCACCT
1861	Qy	TTTGGCAGCTTTGCAACGACATGGTGCCTGGTGGAGAAATGGGTTGGTTCGCCATGTCATT	1920	TTTGGCAGCTTTGCAACGACATGGTGCCTGGTGGAGAAATGGGTTGGTTCGCCATGTCATT
1861	Db	TTTGGCAGCTTTGCAACGACATGGTGCCTGGTGGAGAAATGGGTTGGTTCGCCATGTCATT	1920	TTTGGCAGCTTTGCAACGACATGGTGCCTGGTGGAGAAATGGGTTGGTTCGCCATGTCATT
1921	Qy	GGTTTTTTGTCATTTTTTGCTTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980	GGTTTTTTGTCATTTTTTGCTTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG
1921	Db	GGTTTTTTGTCATTTTTTGCTTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG	1980	GGTTTTTTGTCATTTTTTGCTTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGCCCTG
1981	Qy	GAGCGTGGTTCTCTGTGAAATATTCGCAAAATTTGAAAAGAAAGCTCCAAATTTCTAGC	2040	GAGCGTGGTTCTCTGTGAAATATTCGCAAAATTTGAAAAGAAAGCTCCAAATTTCTAGC
1981	Db	GAGCGTGGTTCTCTGTGAAATATTCGCAAAATTTGAAAAGAAAGCTCCAAATTTCTAGC	2040	GAGCGTGGTTCTCTGTGAAATATTCGCAAAATTTGAAAAGAAAGCTCCAAATTTCTAGC
2041	Qy	CTGAAAGTAATCATTTTGTCTCTGTGCCCTGCTGGCCTTGACCAATGGCCGAGTTCCCTCG	2100	CTGAAAGTAATCATTTTGTCTCTGTGCCCTGCTGGCCTTGACCAATGGCCGAGTTCCCTCG
2041	Db	CTGAAAGTAATCATTTTGTCTCTGTGCCCTGCTGGCCTTGACCAATGGCCGAGTTCCCTCG	2100	CTGAAAGTAATCATTTTGTCTCTGTGCCCTGCTGGCCTTGACCAATGGCCGAGTTCCCTCG
2101	Qy	CTGGGTGGCAGCAATATGGGGCTCCCTCTCTGCTGCTTGGCCCTTTGGGAGCCCTG	2160	CTGGGTGGCAGCAATATGGGGCTCCCTCTCTGCTGCTTGGCCCTTTGGGAGCCCTG
2101	Db	CTGGGTGGCAGCAATATGGGGCTCCCTCTCTGCTGCTTGGCCCTTTGGGAGCCCTG	2160	CTGGGTGGCAGCAATATGGGGCTCCCTCTCTGCTGCTTGGCCCTTTGGGAGCCCTG
2161	Qy	AGCACCATGGCTACATGGTGCCTCATCTTGTCTCAATTCCTTTGCTTCTCATCATG	2220	AGCACCATGGCTACATGGTGCCTCATCTTGTCTCAATTCCTTTGCTTCTCATCATG
2161	Db	AGCACCATGGCTACATGGTGCCTCATCTTGTCTCAATTCCTTTGCTTCTCATCATG	2220	AGCACCATGGCTACATGGTGCCTCATCTTGTCTCAATTCCTTTGCTTCTCATCATG
2221	Qy	ACCAATGGCTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATATTTGG	2280	ACCAATGGCTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATATTTGG
2221	Db	ACCAATGGCTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATATTTGG	2280	ACCAATGGCTACACCAAGCTCTACTGCAATTTGGAACAGGGAGACCTGGAGAAATATTTGG
2281	Qy	GACTGCTCTATGGTAAAAACACATTTGGCCCTGTGTCTTCCACCAACTGCATCTTAACTGC	2340	GACTGCTCTATGGTAAAAACACATTTGGCCCTGTGTCTTCCACCAACTGCATCTTAACTGC
2281	Db	GACTGCTCTATGGTAAAAACACATTTGGCCCTGTGTCTTCCACCAACTGCATCTTAACTGC	2340	GACTGCTCTATGGTAAAAACACATTTGGCCCTGTGTCTTCCACCAACTGCATCTTAACTGC
2341	Qy	CCTGTGGCTTTCTGTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCCCTGAAGTA	2400	CCTGTGGCTTTCTGTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCCCTGAAGTA
2341	Db	CCTGTGGCTTTCTGTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCCCTGAAGTA	2400	CCTGTGGCTTTCTGTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCCCTGAAGTA
2401	Qy	ATTAAGTTTATCTTCTGTGGTGTAGTCCCACTTCCCTGCATGTCTCAATCCCTTCTCTAC	2460	ATTAAGTTTATCTTCTGTGGTGTAGTCCCACTTCCCTGCATGTCTCAATCCCTTCTCTAC

[illegible]

Qy	541	GCCTTTTGAAGCTTTATCGGCATTCGAGCCATGACCTTGGCCCTGAAACAAATACACCAC	600
Ds	589	GCCTTTTGAAGCTTTATCGGCATTCGAGCCATGACCTTGGCCCTGAAACAAATACACCAC	648
Qy	601	ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT	660
Ds	649	ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT	708
Qy	661	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	720
Ds	709	AGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT	768
Qy	721	TTAAATTTACATAATACCTTTGATGAATTCCTCCACTGCAATTAGGACACTCTCCAACTTAAA	780
Ds	769	TTAAATTTACATAATACCTTTGATGAATTCCTCCACTGCAATTAGGACACTCTCCAACTTAAA	828
Qy	781	GAACTAGGATTTTCATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC	840
Ds	829	GAACTAGGATTTTCATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC	888
Qy	841	CCCTCTCTTATACATAATCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT	900
Ds	889	CCCTCTCTTATACATAATCTTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT	948
Qy	901	TTTCAACAATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCACAAATAACTGAA	960
Ds	949	TTTCAACAATTTTACCTGAACTAAGAACACTGACTCTGAAATGGTGCCTCACAAATAACTGAA	1008
Qy	961	TTTCTGATTTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGCACAGATC	1020
Ds	1009	TTTCTGATTTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGCACAGATC	1068
Qy	1021	TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCAGAGTCTGATCTGTCT	1080
Ds	1069	TCATCTCTTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCAGAGTCTGATCTGTCT	1128
Qy	1081	TACAACTTATTAGAAGATTTTACCCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC	1140
Ds	1129	TACAACTTATTAGAAGATTTTACCCAGATTTTTCAGTCTGCCAAAAGCTTCAGAAAATTTGAC	1188
Qy	1141	CTAAGACATATGAATCTACGAAATTAAGTTTGACACTTTTCCAGCAGTTGCTTTAGCCTC	1200
Ds	1189	CTAAGACATATGAATCTACGAAATTAAGTTTGACACTTTTCCAGCAGTTGCTTTAGCCTC	1248
Qy	1201	CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT	1260
Ds	1249	CGATCGCTGAATTTGGCTTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT	1308
Qy	1261	TTGCCATCCCTTAATAAAGCTGGACTATCGTCCAACTCTGCTGCTTTTCCCTATAACT	1320
Ds	1309	TTGCCATCCCTTAATAAAGCTGGACTATCGTCCAACTCTGCTGCTTTTCCCTATAACT	1368
Qy	1321	GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Ds	1369	GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA	1428
Qy	1381	TCATCTGAAACTTTTCCGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAGTGTGT	1440
Ds	1429	TCATCTGAAACTTTTCCGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAGTGTGT	1488
Qy	1441	GCAATTTGGAGTGTGTAGAAATGCCCTTATAAGATTTTCAATCAATGAAATTAAGGTTGACAAAC	1500
Ds	1489	GCAATTTGGAGTGTGTAGAAATGCCCTTATAAGATTTTCAATCAATGAAATTAAGGTTGACAAAC	1548
Qy	1501	AGCAGTATGACGACCTTTTATAGAAAGAGATGCTGGAATTTTTCAGGCTCAAGATGAACGT	1560
Ds	1549	AGCAGTATGACGACCTTTTATAGAAAGAGATGCTGGAATTTTTCAGGCTCAAGATGAACGT	1608
Qy	1561	GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAAGACCTGAAAGCCCTTCAATTCAGTGCAG	1620
Ds	1609	GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAAGACCTGAAAGCCCTTCAATTCAGTGCAG	1668

Qy	1621	TGTTTCACTTCCCAGGCCCTCTTCAAAACCTGTGAAACACCTGTGTATGGCTGGCTGATC	1680
Ds	1669	TGTTTCACTTCCCAGGCCCTCTTCAAAACCTGTGAAACACCTGTGTATGGCTGGCTGATC	1728
Qy	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTATATGCTTTGGTGACTTCA	1740
Ds	1729	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTATATGCTTTGGTGACTTCA	1788
Qy	1741	ACAGTTTTTCAGATCCCTCTGTATCAATTTCCCACTTAACTGTTAAATTTGGGGTCAATCGCA	1800
Ds	1789	ACAGTTTTTCAGATCCCTCTGTATCAATTTCCCACTTAACTGTTAAATTTGGGGTCAATCGCA	1848
Qy	1801	GCAGTGAACATGCTTCAAGGAGTCTCCAGTCCCGTGTGCTGGTGTGGATGCGTTCACT	1860
Ds	1849	GCAGTGAACATGCTTCAAGGAGTCTCCAGTCCCGTGTGCTGGTGTGGATGCGTTCACT	1908
Qy	1861	TTTGGCAGCTTTTGCACACATGCTGTGGGAGAAATGGGGTGGTGGCATGTCAAT	1920
Ds	1909	TTTGGCAGCTTTTGCACACATGCTGTGGGAGAAATGGGGTGGTGGCATGTCAAT	1968
Qy	1921	GGTTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTTCTGCTTACTCTGCGCAGCCCTG	1980
Ds	1969	GGTTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTTCTGCTTACTCTGCGCAGCCCTG	2028
Qy	1981	GAGCGTGGGTTCTCTGTGAAAATATTTCTGCAAAATTTTGAACGAAAGCTTCAATTTTCTAGC	2040
Ds	2029	GAGCGTGGGTTCTCTGTGAAAATATTTCTGCAAAATTTTGAACGAAAGCTTCAATTTTCTAGC	2088
Qy	2041	CTGAAAGTAAATCAATTTTGTCTGTGCGCTGTGCGCTTGACCATGGCGGCGAGTTCCCTCTG	2100
Ds	2089	CTGAAAGTAAATCAATTTTGTCTGTGCGCTGTGCGCTTGACCATGGCGGCGAGTTCCCTCTG	2148
Qy	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGGGGAGGCC	2160
Ds	2149	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTTGGGGAGGCC	2208
Qy	2161	AGCACCATGGGCTACATGGTGGCTCTCATTTGCTCAAATTTCCCTTTGCTTCCCTCATGATG	2220
Ds	2209	AGCACCATGGGCTACATGGTGGCTCTCATTTGCTCAAATTTCCCTTTGCTTCCCTCATGATG	2268
Qy	2221	ACCATGGCTTACACCAAGCTCTACTGCAATTTTGACAAAGGAGACCTGGAGAAATTTTGG	2280
Ds	2269	ACCATGGCTTACACCAAGCTCTACTGCAATTTTGACAAAGGAGACCTGGAGAAATTTTGG	2328
Qy	2281	GACTGCTTATGTGTAAACACATTTCCCTGTCTTTTACCAACTGCATCTTAAACTGC	2340
Ds	2329	GACTGCTTATGTGTAAACACATTTCCCTGTCTTTTACCAACTGCATCTTAAACTGC	2388
Qy	2341	CCTGTGGCTTTTCTTGTCTTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAAGTA	2400
Ds	2389	CCTGTGGCTTTTCTTGTCTTCTCTCTTAAATAAACCTTACATTTATCAGTCTCTGAAAGTA	2448
Qy	2401	ATTAAGTTTATCTCTTCTGGTGGTAGTCCCATCTTCTGCAATGCTCAATCCCTTCTCTAC	2460
Ds	2449	ATTAAGTTTATCTCTTCTGGTGGTAGTCCCATCTTCTGCAATGCTCAATCCCTTCTCTAC	2508
Qy	2461	ATCTTGTTCATCTCTCACTTTTAAAGGAGATCTGGTGGAGCTTGAGAAAGCAAACTACGTC	2520
Ds	2509	ATCTTGTTCATCTCTCACTTTTAAAGGAGATCTGGTGGAGCTTGAGAAAGCAAACTACGTC	2568
Qy	2521	TGGAACAAGATCAAAACACCCCAAGCTTGTATGATGCTTAACTCTGATGATGTCGAAAAACAG	2580
Ds	2569	TGGAACAAGATCAAAACACCCCAAGCTTGTATGATGCTTAACTCTGATGATGTCGAAAAACAG	2628
Qy	2581	TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCATTATGACCTG	2640
Ds	2629	TCCTGTGACTCAACTCAAGCCTTGGTAACTTTTACAGCTCCAGCATCATTATGACCTG	2688
Qy	2641	CCTCCCAAGTTTCCGTGGCCATCACAGCTTATCCAGTGTAGAGCTGCATCTTTTCTCTCT	2700
Ds	2689	CCTCCCAAGTTTCCGTGGCCATCACAGCTTATCCAGTGTAGAGCTGCATCTTTTCTCTCT	2748
Qy	2701	GTGGCATTTGTCCCATGTCTCTAA	2724

QY	1261	TTGGCATCCCTAATAAAGCTGGACCTATCGTGCCAACTCCTGTCGTCTTTTCTATAACT	1320
Db	1309	TTGCCATCCCTATAAAGCTGGACCTATCGTGCCAACTCCTGTCGTCTTTTCTATAACT	1368
QY	1321	GGGTTACATGGTTTAACTCACTTAAATAATAACAGGAAATCATGCTTACAGAGCTTGATA	1380
Db	1369	GGGTTACATGGTTTAACTCACTTAAATAATAACAGGAAATCATGCTTACAGAGCTTGATA	1428
QY	1381	TCATCTGAAACCTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGTCTGT	1440
Db	1429	TCATCTGAAACCTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGTCTGT	1488
QY	1441	GCATTTGGAGTGTGTGAAGATGCCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC	1500
Db	1489	GCATTTGGAGTGTGTGAAGATGCCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAAC	1548
QY	1501	AGCAGTATGGAACGACTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560
Db	1549	AGCAGTATGGAACGACTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1608
QY	1561	GACCTTGAAAGTTTCCTGCTTGACTTTTGAGGAGACCTGGAAGCCCTTCATTCAAGTGCAG	1620
Db	1609	GACCTTGAAAGTTTCCTGCTTGACTTTTGAGGAGACCTGGAAGCCCTTCATTCAAGTGCAG	1668
QY	1621	TGTTTCACTTCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC	1680
Db	1669	TGTTTCACTTCCCAGAGCCCTTCAAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC	1728
QY	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAATGCTTTGGTGACTTCA	1740
Db	1729	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAATGCTTTGGTGACTTCA	1788
QY	1741	ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTAAATTTGGGTTCATCGCA	1800
Db	1789	ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTAAATTTGGGTTCATCGCA	1848
QY	1801	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGGCTGGTGTGGATGCGTTCACT	1860
Db	1849	GCAGTGAACATGCTCACGGAGTCTCCAGATGCCGTGCTGGCTGGTGTGGATGCGTTCACT	1908
QY	1861	TTTTGGCAGCTTTGCAACACATGGTGTGCTGGTGGAGAAATGGGTTGGTTGCCATGTCAATT	1920
Db	1909	TTTTGGCAGCTTTGCAACACATGGTGTGCTGGTGGAGAAATGGGTTGGTTGCCATGTCAATT	1968
QY	1921	GGTTTTTGTGCCATTTTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGGAGCCCTG	1980
Db	1969	GGTTTTTGTGCCATTTTTTGTCTCAGAAATCATCTGTTTTCTGCTTACTCTGGGAGCCCTG	2028
QY	1981	GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTTGAAACGAAAGCTCCAAATTTCTAGC	2040
Db	2029	GAGCGTGGGTTCTCTGTGAAATATTTCTGCAAAATTTTGAAACGAAAGCTCCAAATTTCTAGC	2088
QY	2041	CTGAAAGTAATCATTTTTGCTCTGTGCCCTGTGCTGGCTTGACCATGGCGCAGTTTCCCTG	2100
Db	2089	CTGAAAGTAATCATTTTTGCTCTGTGCCCTGTGCTGGCTTGACCATGGCGCAGTTTCCCTG	2148
QY	2101	CTGGGTGGCAGCAAGTATGGGGCCCTCCCTCTCTGGCTGCTTTGGCTTTTGGGGAGCCC	2160
Db	2149	CTGGGTGGCAGCAAGTATGGGGCCCTCCCTCTCTGGCTGCTTTGGCTTTTGGGGAGCCC	2208
QY	2161	AGCACCATGGCTACATGGTGGCTCATCTTGCTCAATTTCCCTTTGCTTCCCTCATCATG	2220
Db	2209	AGCACCATGGCTACATGGTGGCTCATCTTGCTCAATTTCCCTTTGCTTCCCTCATCATG	2268
QY	2221	ACCATTGCCTACACCAAGCTCTACTGCAATTTTGGACAAAGGAGACCTGGAGAAATTTTGG	2280
Db	2269	ACCATTGCCTACACCAAGCTCTACTGCAATTTTGGACAAAGGAGACCTGGAGAAATTTTGG	2328
QY	2281	GACTGCTCTATGGTAAACACATTTGGCCCTGTGCTTCTTCCAACTGCAATCCTTAACTGC	2340
Db	2329	GACTGCTCTATGGTAAACACATTTGGCCCTGTGCTTCTTCCAACTGCAATCCTTAACTGC	2388

[illegible]

QY 301 GCTCTGACATACATTTCCCAAGGAGCAATTCACCTGGCCCTTTACAGTCTTTAAAGTTCTTTATG 360
Db |||||
QY 301 GCTCTGACATACATTTCCCAAGGAGCAATTCACCTGGCCCTTTACAGTCTTTAAAGTTCTTTATG 360
Db |||||
QY 361 CTGCAGAAATAATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTTCGGAAGCCCTT 420
Db |||||
QY 361 CTGCAGAAATAATCAGCTAAGACACAGTACCCACAGAGCTCTGCAGAAATTTTCGGAAGCCCTT 420
Db |||||
QY 421 CAATCCCTGCGTCTGGATGCTAAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db |||||
QY 421 CAATCCCTGCGTCTGGATGCTAAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db |||||
QY 481 CTGCATTTCCCTGAGGACCTCTGGCTGGATGCAATGCTTTAAACAGAAATCCCGTCCAG 540
Db |||||
QY 481 CTGCATTTCCCTGAGGACCTCTGGCTGGATGCAATGCTTTAAACAGAAATCCCGTCCAG 540
Db |||||
QY 541 GCTTTTGAAGTTTATCGGCAATTCGAAGCCATGACCTTTGGCCCTGAAACAAATAACACCAC 600
Db |||||
QY 541 GCTTTTGAAGTTTATCGGCAATTCGAAGCCATGACCTTTGGCCCTGAAACAAATAACACCAC 600
Db |||||
QY 601 ATACAGACTATGCTTTGGAAAACCTCTCAGCTTGTAGTTCTACATCTCCATAACAAAT 660
Db |||||
QY 601 ATACAGACTATGCTTTGGAAAACCTCTCAGCTTGTAGTTCTACATCTCCATAACAAAT 660
Db |||||
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db |||||
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db |||||
QY 721 TTAATAATCAATAACCTTTGATGAATTCGCCACTGCAATTAGGACACTCTCCAACTTTAAA 780
Db |||||
QY 721 TTAATAATCAATAACCTTTGATGAATTCGCCACTGCAATTAGGACACTCTCCAACTTTAAA 780
Db |||||
QY 781 GAACTAGAGATTCATAGCAACAATATCAGGTGCGATACCTAGAGAAAGCAATTTGTAGGCAAC 840
Db |||||
QY 781 GAACTAGAGATTCATAGCAACAATATCAGGTGCGATACCTAGAGAAAGCAATTTGTAGGCAAC 840
Db |||||
QY 841 CCTTCTCTTATAGCAATACATTTCTATGACAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db |||||
QY 841 CCTTCTCTTATAGCAATACATTTCTATGACAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db |||||
QY 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTCAACAATAACTGAA 960
Db |||||
QY 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTCAACAATAACTGAA 960
Db |||||
QY 961 TTTCTGATTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGGACAGATC 1020
Db |||||
QY 961 TTTCTGATTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGGACAGATC 1020
Db |||||
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
Db |||||
QY 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
Db |||||
QY 1081 TACAACCTTATTAGAAGATTTACCCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db |||||
QY 1081 TACAACCTTATTAGAAGATTTACCCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
Db |||||
QY 1141 CTAAGACATAATGAAATCTAGAAATTTAAAGTTGACACTTTCCAGCAGTTCGTTAGCCCTC 1200
Db |||||
QY 1141 CTAAGACATAATGAAATCTAGAAATTTAAAGTTGACACTTTCCAGCAGTTCGTTAGCCCTC 1200
Db |||||
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCGATTTATTCACCCCAATGCAATTTTCCAAT 1260
Db |||||
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTCGATTTATTCACCCCAATGCAATTTTCCAAT 1260
Db |||||
QY 1261 TTGCCATCCCTTAATAAGCTGGAACCTATCGTCCAACTCTGTGCTGCTTTTCTTATAACT 1320
Db |||||
QY 1261 TTGCCATCCCTTAATAAGCTGGAACCTATCGTCCAACTCTGTGCTGCTTTTCTTATAACT 1320
Db |||||
QY 1321 GGGTTACATGGTTTAACTCATTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db |||||
QY 1321 GGGTTACATGGTTTAACTCATTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
Db |||||
QY 1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTACCAGTCTGT 1440
Db |||||

Db |||||
QY 1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCTTATGCTTACCAGTCTGT 1440
Db |||||
QY 1441 GCATTTGAGTGTGTGAGAAATCCCTATAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
Db |||||
QY 1441 GCATTTGAGTGTGTGAGAAATCCCTATAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
Db |||||
QY 1501 AGCAGTATGAGACGACTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db |||||
QY 1501 AGCAGTATGAGACGACTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db |||||
QY 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db |||||
QY 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db |||||
QY 1621 TGTTCACCTTCCAGGCCCCCTTCAAAACCTGTGGAACACCTGCTGTGATGGCTGCTGATC 1680
Db |||||
QY 1621 TGTTCACCTTCCAGGCCCCCTTCAAAACCTGTGGAACACCTGCTGTGATGGCTGCTGATC 1680
Db |||||
QY 1681 AGAATTCGAGTGTGGAACCATAGCAGTTCTGSCACTTACTTCTAATGCTTTTGGTGCATTTCA 1740
Db |||||
QY 1681 AGAATTCGAGTGTGGAACCATAGCAGTTCTGSCACTTACTTCTAATGCTTTTGGTGCATTTCA 1740
Db |||||
QY 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTTAAACTGTTAAATTTGGGGTCAATGCA 1800
Db |||||
QY 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTTAAACTGTTAAATTTGGGGTCAATGCA 1800
Db |||||
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCGCTGTGGTGTGATGGTGCCTTCACT 1860
Db |||||
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCGCTGTGGTGTGATGGTGCCTTCACT 1860
Db |||||
QY 1861 TTTGCGAGCTTTGACACGACATGCTGCTGGTGGGAGAAATGGGGTGTGGTGCATGTCAAT 1920
Db |||||
QY 1861 TTTGCGAGCTTTGACACGACATGCTGCTGGTGGGAGAAATGGGGTGTGGTGCATGTCAAT 1920
Db |||||
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTTCTGCTTACTCTCGGAGCCCTG 1980
Db |||||
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTTCTGCTTACTCTCGGAGCCCTG 1980
Db |||||
QY 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTTGAAGCAAAAGCTCCCAATTTCTAGC 2040
Db |||||
QY 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTTGAAGCAAAAGCTCCCAATTTCTAGC 2040
Db |||||
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGACCATGGCGCAGTTCCTCCCTG 2100
Db |||||
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGACCATGGCGCAGTTCCTCCCTG 2100
Db |||||
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTTTTGGGGAGGCC 2160
Db |||||
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTTTTGGGGAGGCC 2160
Db |||||
QY 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTCCTTCTCATGATG 2220
Db |||||
QY 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTCCTTCTCATGATG 2220
Db |||||
QY 2221 ACCAATTCCTACACCAAGCTCTACTGCAATTTGACCAAGGAGAGACCTGGAGAAATTTGG 2280
Db |||||
QY 2221 ACCAATTCCTACACCAAGCTCTACTGCAATTTGACCAAGGAGAGACCTGGAGAAATTTGG 2280
Db |||||
QY 2281 GACTGCTCTATGTTGTAACCAATTCATGCTGCTTCTTCAACCACTGCATCCTTAAACTGTC 2340
Db |||||
QY 2281 GACTGCTCTATGTTGTAACCAATTCATGCTGCTTCTTCAACCACTGCATCCTTAAACTGTC 2340
Db |||||
QY 2341 CCTGTGGCTTCTTGTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
Db |||||
QY 2341 CCTGTGGCTTCTTGTCTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2400
Db |||||
QY 2401 ATTAAGTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2460
Db |||||
QY 2401 ATTAAGTTTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2460
Db |||||
QY 2461 ATCTTGTTCATCTCTCTTTCAGGAGGATCTGTGAGCCTGTGAGAAAGCAAACTTACGTC 2520
Db |||||

Qy	961	TTTCTGATTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGCACAGATC	1020
Db	961	TTTCTGATTTAACTGGAACTGCAAACTCGAGAGTCTGACTTTAACTGGAGCACAGATC	1020
Qy	1021	TCATCTCTTCTGCTAAACCGGTCTGCAATCAGTTACCTTAATCTCCAAGTGTCTAGATCTGTCT	1080
Db	1021	TCATCTCTTCTGCTAAACCGGTCTGCAATCAGTTACCTTAATCTCCAAGTGTCTAGATCTGTCT	1080
Qy	1081	TACAACCTATTAGAAAGATTACCCAGGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC	1140
Db	1081	TACAACCTATTAGAAAGATTACCCAGGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTGAC	1140
Qy	1141	CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC	1200
Db	1141	CTAAGACATAATGAAATCTACGAAATTAAGTTGACACTTTCCAGCAGTTGCTTAGCCCTC	1200
Qy	1201	CGATCGCTGAAATTTGGCTTGGAACAAAATTTGCTATTATTCCACCCCAATGCAATTTCCACT	1260
Db	1201	CGATCGCTGAAATTTGGCTTGGAACAAAATTTGCTATTATTCCACCCCAATGCAATTTCCACT	1260
Qy	1261	TTGCGATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCTCTATAAAT	1320
Db	1261	TTGCGATCCCTTAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCTCTATAAAT	1320
Qy	1321	GGGTTACATGGTTTTAACTCACTTAAAAATTAAACAGGAATCATGCCCTTACAGAGCTTGAATA	1380
Db	1321	GGGTTACATGGTTTTAACTCACTTAAAAATTAAACAGGAATCATGCCCTTACAGAGCTTGAATA	1380
Qy	1381	TCATCTGAAAACCTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAAGTGTGTGT	1440
Db	1381	TCATCTGAAAACCTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAAGTGTGTGT	1440
Qy	1441	GCATTTGGAGTGTGTGAGATGCTTAAAGATTTCTTAATCAATGGGAATAAAGGTGACAAC	1500
Db	1441	GCATTTGGAGTGTGTGAGATGCTTAAAGATTTCTTAATCAATGGGAATAAAGGTGACAAC	1500
Qy	1501	AGCAGTATGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
Db	1501	AGCAGTATGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
Qy	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
Db	1561	GACCTTGAAGATTTCTGCTTGACTTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGCAG	1620
Qy	1621	TGTTCACTTCCACAGGCCCTTCAAACTCTGTGAACACCTGCTTGATGCTCGCTGATC	1680
Db	1621	TGTTCACTTCCACAGGCCCTTCAAACTCTGTGAACACCTGCTTGATGCTCGCTGATC	1680
Qy	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAATGCTTTGGTGCATCTCA	1740
Db	1681	AGAAATGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAATGCTTTGGTGCATCTCA	1740
Qy	1741	ACAGTTTTAGATCCCTCTGTACATTTCCCCCAATTAAATCTGTTAATTTGGGTGATCGCA	1800
Db	1741	ACAGTTTTAGATCCCTCTGTACATTTCCCCCAATTAAATCTGTTAATTTGGGTGATCGCA	1800
Qy	1801	GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGCTGGTGTGGATGGTTCCACT	1860
Db	1801	GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGCTGCTGGTGTGGATGGTTCCACT	1860
Qy	1861	TTTGGCAGCTTTGCAAGACATGGTGCCTGGTGGAGAAATGGGGTTGGTGTGCCATGTCAATT	1920
Db	1861	TTTGGCAGCTTTGCAAGACATGGTGCCTGGTGGAGAAATGGGGTTGGTGTGCCATGTCAATT	1920
Qy	1921	GGTTTTTTTGTCCATTTTTTGTCTCAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG	1980
Db	1921	GGTTTTTTTGTCCATTTTTTGTCTCAGAAATCATCTGTTTTTCTGCTTACTCTGGCAGCCCTG	1980
Qy	1981	GAGCGTGGGTCTCTGTGGAATATTCTGCAAAATTTGAAACGAAAGCTCCAATTTCTAGC	2040
Db	1981	GAGCGTGGGTCTCTGTGGAATATTCTGCAAAATTTGAAACGAAAGCTCCAATTTCTAGC	2040
Qy	2041	CTGAAAGTAATCAATTTTTGCTCTGTGCCCTCTGCTGCCCTTTGACCAATGGCGCAGTTCCCTCTG	2100

Db	2041	CTGAAAGTAATCATTTTGGCTCTGTGCGCTGTGCGCTTGACCAATGGCGCAGTTCCCGCTG	2100
Qy	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGTGCTTTGGCTTTGGGGAGCCCC	2160
Db	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTTGTGCTTTGGCTTTGGGGAGCCCC	2160
Qy	2161	AGCACCATGGGCTACATGGTTCGCTCTCATCTTGCTCAAAATTCCTTTGCTTCCTCATGATG	2220
Db	2161	AGCACCATGGGCTACATGGTTCGCTCTCATCTTGCTCAAAATTCCTTTGCTTCCTCATGATG	2220
Qy	2221	ACCAATTCGCTACACCAAGCTCTATGCAATTTGGACAAGGGAGACCTGGAGAAATATTTGG	2280
Db	2221	ACCAATTCGCTACACCAAGCTCTATGCAATTTGGACAAGGGAGACCTGGAGAAATATTTGG	2280
Qy	2281	GACTGCTCTATGTAAGAACACATTTGGCGCTGCTCTTTCACCAACTGCATCTCTAAACTGC	2340
Db	2281	GACTGCTCTATGTAAGAACACATTTGGCGCTGCTCTTTCACCAACTGCATCTCTAAACTGC	2340
Qy	2341	CCTGTGGCTTTCTTGCTCTCTCTCTCTTTTAATAAAACCTTACATTTATCAGTCTCTGAAGTA	2400
Db	2341	CCTGTGGCTTTCTTGCTCTCTCTCTCTTTTAATAAAACCTTACATTTATCAGTCTCTGAAGTA	2400
Qy	2401	ATTAAGTTTATCCTTCTGTGGTAGTCCCACTCTCTGCGATGTCTCAATCCCTTCTCTAC	2460
Db	2401	ATTAAGTTTATCCTTCTGTGGTAGTCCCACTCTCTGCGATGTCTCAATCCCTTCTCTAC	2460
Qy	2461	ATCTTGTTCAATCCTCTCACTTTTAAGGAGGATCTGCTGAGCCTGAGAAAGCAAACTACGTC	2520
Db	2461	ATCTTGTTCAATCCTCTCACTTTTAAGGAGGATCTGCTGAGCCTGAGAAAGCAAACTACGTC	2520
Qy	2521	TGGACAAGATCAAAAACACCCCAAGCTTGATGTCAATTAACCTCTGATGATGTGCGAAAAACAG	2580
Db	2521	TGGACAAGATCAAAAACACCCCAAGCTTGATGTCAATTAACCTCTGATGATGTGCGAAAAACAG	2580
Qy	2581	TCTGTGACTCAACTCAAGCCTTGGTAAACCTTTACAGCTCCAGCATCACTTATGACCTG	2640
Db	2581	TCTGTGACTCAACTCAAGCCTTGGTAAACCTTTACAGCTCCAGCATCACTTATGACCTG	2640
Qy	2641	CTCTCCAGTTTCCGTGCCATCACCAGCTTATCCAGTGCATGAGAGCTGCCATCTTCTCTCT	2700
Db	2641	CTCTCCAGTTTCCGTGCCATCACCAGCTTATCCAGTGCATGAGAGCTGCCATCTTCTCTCT	2700
Qy	2701	GTGGCATTTGTCGCCATGCTCTCTAA 2724	
Db	2701	GTGGCATTTGTCGCCATGCTCTCTAA 2724	
RESULT 11			
BD135244			
LOCUS	BD135244	2082 bp DNA linear	PAT 18-SEP-2002
DEFINITION	Novel mammalian G protein-coupled receptor having extracellular leucine-rich repeating domain.		
ACCESSION	BD135244		
VERSION	BD135244.1	GI:23230189	
KEYWORDS	JP 2002507406-A/2.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 2082)		
TITLE	Haueh,A.J.W., Hsu,S.Y., Liang,S.G. and Spe,P.J.V.D. Novel mammalian G protein-coupled receptor having extracellular leucine-rich repeating domain		
JOURNAL	Patent: JP 2002507406-A 2 12-MAR-2002; THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY, AKZO NOBEL NV		
COMMENT	OS Homo sapiens (human)		
	PN JP 2002507406-A/2		
	PD 12-MAR-2002		
	PF 25-MAR-1999 JP 2000537903		
	PI 26-MAR-1998 US 60/079501		
	PR AARON J W HSUEH, SHEAU YU HSU, SHAN GUANG LIANG,PETRUS JOHANNES		

Db 1861 AGAAGCAAACTACGCTGACAGAGTCAAAAACCCAGCTTGAATGCTCAATTAATCT 1920
Qy 2563 GATGATGCGAAACAGTCTCTGAGTCAACTCAAGCTTTGGTAACCTTTACAGCTCC 2622
Db 1921 GATGATGCGAAACAGTCTCTGAGTCAACTCAAGCTTTGGTAACCTTTACAGCTCC 1980
Qy 2623 AGCATCACTATGACCTGCTCCAGTTCCTGCGCATCACCAGCTTATCCAGTACTGAG 2682
Db 1981 AGCATCACTATGACCTGCTCCAGTTCCTGCGCATCACCAGCTTATCCAGTACTGAG 2040
Qy 2683 AGCTGCCATCTTCTCTGTCGCAATTTGTCCTCCATGCTCTCTAA 2724
Db 2041 AGCTGCCATCTTCTCTGTCGCAATTTGTCCTCCATGCTCTCTAA 2082

RESULT 12
AF110818 3115 bp mRNA linear ROD 11-FEB-1999
LOCUS Mus musculus orphan G protein-coupled receptor FEX mRNA, complete
DEFINITION cds.

ACCESSION AF110818
VERSION AF110818.1 GI:4262545
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3115)

REFERENCE
Hermey, G., Methner, A., Schaller, H.C. and Hermans-Borgmeyer, I.
AUTHORS Identification of a novel seven-transmembrane receptor with
TITLE homology to glycoprotein receptors and its expression in the adult
and developing mouse

JOURNAL Biochem. Biophys. Res. Commun. 254 (1), 273-279 (1999)
MEDLINE 99121227
PUBMED 9920770

REFERENCE
2 (bases 1 to 3115)
Hermey, G., Methner, A., Schaller, H.C. and Hermans-Borgmeyer, I.
AUTHORS Direct Submission
TITLE Submitted (03-DEC-1998) Center for Molecular Neurobiology,
JOURNAL Martinistr. 52, Hamburg D-20246, Germany
FEATURES Location/Qualifiers

source
1..3115
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"

CDS
1..2724
/note="contains leucine-rich repeats"

/codon_start=1
/product="orphan G protein-coupled receptor FEX"
/protein_id="AADI4684.1"

/db_xref="GI:4262546"
/translation="MDTSCVHMLLSLLALLQVAAGSSPGDAIPRGCPSHCHCELDG
RMLRVDCSLGLSELNSLVSFTSYLDLSMNNISOLPASLHLRLCFLEELRLAGNAL
THIPKGAFTGLKVLMLNNQRLKVPKEALQNLRSIQSLRLDANHSVYPPSCFSG
LHSRLHLDNALITDVPQAFRSIALQAMTLANKIHLIADYAFGNLSLVVHLH
NNRILHSGLKCFDGLSLNLDLDEFITAKLNSLDELKELGFSNNIRSIPEBA
FVGNPSLITHTFYDNPTQFVGSFAQLPELRLTLINGASHITFPFLHTGTATLESIT
LTGAKISLPQAVCDQNLQVLSYLLDEPLSLGCGQLKQIDLRHNEIYKGS
TFQOLFNLRSANPELKIITHPNAPSLPSLILKLDLSNLLSPFPVTVGLHGLTHLKL
TGNEALQSLIPANPELKIITEMPAYVCCAFGCGENVYKISNOWNKDDGNSVDDLHK
KDAGLFQVQDERDLEDFLLDEEDLNALHSVQSPSPGPKPCHEHLPGLSLRLIGVNT
TAVLTSNCNALVALVTFRPLTIISSIKLLIGVIAVDILMGVSSAVLAADVAFTRFR
AQHGWEDGQIQVGFSLIFASESSIFLLTLAALRGFSVKCSKFEKVPKAPLFSR
AIVLLCVLLATITATPLLAGSKYNAPCLPLPFGPSTTGYNVALVLLNSLCFLIM
TIATYKLYCSKELELWDCSMVKHIALLLFANCLICPVAFVLSFLINLTFISF
DVIKFILLIVLPCLNPLIYIVFNPHFKDMSLGHKTRFWMRSKHASILLNSDD
VEKRSCESTQALVSFTHASIAYDLPSISGASPAYPMTESCHLSVAVFVPCFL"

ORIGIN

Query Match 75.4%; Score 2052.6; DB 10; Length 3115;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 2304; Conservative 0; Mismatches 419; Indels 0; Gaps 0;

Qy 1 ATGGACACCTCCCGCTCGGTGCTCTCTGCTCCCTTGGCTGTGCTGCAGCTGCGACGACC 60
Db 1 ATGGACACCTCCCTCGTCCATCATGCTCTGCTGCTGCTGCGGCTGCTGCACTTGTGGCC 60
Qy 61 GGGGCGAGCTCTCCAGGCTGTGGTGTGCTGAGGGGCTGCCCCACACACACTGTCAATGC 120
Db 61 GCGGCGAGCTCACCGGACAGATGCGATACCGGGGGCTGCCCATCACACTGTCACTGT 120
Qy 121 GAGCCGACGCGAGGATGTTGCTCAGGGTGAATGCTCCGACCTGGGGCTCTCGAGCTG 180
Db 121 GAGCTGATGCGAGGATGCTCTCAGGGTGAATGCTCGGACCTGGGGCTCTCGAGCTG 180
Qy 181 CCTTCCAACTCAGCTCTTCACTCTACCTCTACCTCTGAGCTCTAGTATGAACAACATCAGTCAG 240
Db 181 CCTTCCAACTCAGCTCTTCACTCTACCTCTGAGCTCTAGTATGAACAACATCAGTCAG 240
Qy 241 CTGCTCCCGAATCCCTGCGCAGTCTCGCTTCTCGGAGGAGTACGCTTTGCGGGAAC 300
Db 241 CTACCCGCGAGTCTCTACATCGCTCTGCTTCTAGAGAGTACGCTTGTGCTGGAAT 300
Qy 301 GCTCTGACATACATTTCCAAAGGAGGACATTCATCGGCTTTTACAGTCTTAAAGTCTTATG 360
Db 301 GCTTTGACACATTTCCAAAGGAGGAGCTTCAAGGGCTTTCAGGCTCTCAAGCTCAAAAGTCTTATG 360
Qy 361 CTGCAGATTAATCAGCTAAGACACGTACCCACACAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db 361 CTGCAGAAACACAGCTGAGAAAGGTTCCGGAGGAAGCGCTACAGAAATTTGAGAAGCCTT 420
Qy 421 CAATCCCTGCGTCTGGATGTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGGC 480
Db 421 CAATCCCTGCGCTAGATGCCAACACATCAGCTACGTGCCACCAGCTGTTTCAAGGCGC 480
Qy 481 CTGCAATTCCTGAGGACCTGTGTGCTGATGACAATGCTCTCAGAGAGCTCCCTGTCTCAG 540
Db 481 CTGCACTCCCTGAGGACCTGTGTGCTAGATGACAATGCTCTCAGAGAGCTCCCTGTCTCAG 540
Qy 541 GCTTTTGAAGTATTTCGCAATTCGAGGACATGCTTGGCCCTGAGCAAAATACACAC 600
Db 541 GCTTTTGAAGTATTTCAGCCCTGCAAGCCATGACCTTGGCCCTGAGCAAAATACACAC 600
Qy 601 ATACGAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTCTTACATCTCCATAACAAT 660
Db 601 ATACGAGACTACGCTTTGGAAACCTCTCCAGCTCTGTTGCTTCTGCATCTCCATAAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGACTCCAAGCTTGGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGACTCCAAGCTTGGAGACTTTAGAT 720
Qy 721 TTAATTTACATAACCTTGTATGTAATTCCTCACTGCAATTAGGACACTCTCCAACTTAAA 780
Db 721 TTAATTTATATAACCTTGTATGTAATTCCTCACTGCAATTAAGACACTCTCCAACTTAA 780
Qy 781 GAACTAGGATTTCAATAGCAACAATATCAGGTCGATACCTCGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTCCACAGCAACAATATCAGGTCATACCGGAGCGAGCGTTCGTAGGCAAC 840
Qy 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCAAATTTGTTGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATACATTTCTATGACAAACCCCATCCAAATTTGTTGGAGATCTGCT 900
Qy 901 TTTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGGTGGCTCAACAATAACTGAA 960
Db 901 TTTTCAACATTTTACCTGAACCTAAGAACACTGACTCTGAATGGTGGCTCAACAATAACTGAA 960
Qy 961 TTTTCTGATTTAACTGGAACCTGCAAACTCGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Db 961 TTTTCTGATTTAACTGGAACCTGCAAACTCGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Qy 1021 TCATCTCTTCTCTAAAACCGTCTGCAATCAGTTTACCTTAATCTCCAAAGTCTAGATCTGTCT 1080
Db 1021 TCATCTCTTCTCTAAAACCGTCTGATCAGTTTACCTTAATCTCCAAAGTCTAGATTTGTCT 1080

1081 QY TACAACCTATTAGAGATTATCCAGTCTTTTTCAGTCTGCCAAAGCTTCAGAAATTTGAC 1140
1081 Db TACAACCTACTAGCAGACTTACCCAGTCTTTTGTGAGGCTGCCAAAGCTTCAGAAATTTGAC 1140
1141 QY CTAAGACATAAATGAAATCTACGAAATTAAGTTTGCACATTTCCAGCAGTGTCTTACGCTC 1200
1141 Db CTGAGGCATACGAGATCTATGAAATTAAGGCGAGCATTTTTCAGCAGTGTCTTACGCTC 1200
1201 QY CGATCGCTGAATTTGGCTTTGGAACAAAATTTGCTATTATTCACCCCAANGCATTTTCCACT 1260
1201 Db CGATCTCTGAATTTAGCTAGTGAATAAATTTGCTATTATTCACCCCAANGCATTTTCTACG 1260
1261 QY TTGGCATCCCTAATAAGCTGACCTATCGTCAACCTCTCTGCTCTTCTCTTAATCT 1320
1261 Db TTGGCCGCTCTAATAAGTTGACCTATCATCAATCTCTGCTCTCTTCCCTGTGACT 1320
1321 QY GGGTTACATGTTTAACTCACTTAAATTAACAGGAATCATGCTTACAGAGCTTGATA 1380
1321 Db GGGTTACATGTTTAACTCACTTAAATTAACAGGACCGAGCTTACAGAGCTTGATA 1380
1381 QY TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTTACCAGTGTGT 1440
1381 Db CCATCTGCAAACTTCCAGAGCTCAAGATTATAGAAATGCTTATGCTTTACCAGTGTGT 1440
1441 QY GCATTTGGAGTGTGAGAAATGCTTATAGAAATTTCTAATCAATGGAATAAAGGTGACAA 1500
1441 Db GCATTTGGGGGTGTGAGAAATGCTTATAGAAATTTCTAATCAATGGAATAAAGACGCGC 1500
1501 QY AGCAGTATGAGACCTTCAATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAAGT 1560
1501 Db AACAGTGTGAGACCTTCAATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAGCGG 1560
1561 QY GACCTTTGAAGATTTCTGCTTCACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
1561 Db GACCTTTGAAGATTTCTGCTTCACTTTGAGGAAGACCTGAAAGCCCTTCACTCGGTGCAG 1620
1621 QY TGTTCACCTTCCAGGCCCCCTTCAAAACCTGTGAAACCTGCTGTGATGCTGGCTGATC 1680
1621 Db TGTCTGCTTCCAGGCCCCCTTCAAGCCCTGTGAGACCTATTTGGTAGCTGGCTGATC 1680
1681 QY AGAATTTGAGTGTGAGACCTATGAGTCTGAGCATTCTGTAATGCTTTGCTGACTTCA 1740
1681 Db GAAATCGGGGTGTGAGACCTGAGCAGTACTGAGCTTTCTGCAATGCTTTGCTGGCTTTG 1740
1741 QY ACAGTCTTTCAGATCCCTCTGTACATTTCCCCCACTTAAATGTTAAATGGGGTCAATCGCA 1800
1741 Db ACCGTCTTCAGAACTCCCTCTGTACATCTTCCATAAAGCTGCTAATTTGGGTATCGCG 1800
1801 QY CGAGTGAACATGCTCAAGGAGTCTCCAGTCCGCTGCTGGCTGGTGTGATGCTTCACT 1860
1801 Db GTAGTGGACATTTCTCATGGGGTCTCCAGTCTGCTGGCTGGCTGGATGCAATTCAT 1860
1861 QY TTTGGCAGCTTTGACGACATGCTGCTGGGAGATGGGGTGGTGGCCATGCTCAT 1920
1861 Db TTTGGCCGCTTTGCTGACGACGCTGGTGGGAGAGCGAATCGGCTGCCAAATTCGTT 1920
1921 QY GGTCTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAGCCCTG 1980
1921 Db GGTCTTCTGCAATTTTGTCTCCGAAATGCTGCAATCTTCTGCTCACTCTGGCAGGGCTG 1980
1981 QY GAGCGTGGGTTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
1981 Db GAACGAGGTTTTTCTGTCAAGTGTCTTCCAAAGTTTGAAGTGAAGCTCCCTTTTTCAGC 2040
2041 QY CTGAAAGTAACTATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
2041 Db CTGAGAGGCAATGTTTGTGTATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
2101 QY CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
2101 Db CTAGAGGGCAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2160
2161 QY AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220

2161 Db AGCACACGGGCTACATGGTGGCTCTCGTGTGCTCAACTCTCTCTGTTTCTCTATAATG 2220
2221 QY ACCATTGGCTACACCAAGCTCTACTGCAATTTTGGCAAGGAGACCTCGAGAAATATTGG 2280
2221 Db ACCATTGGCTACACCAAGCTCTACTGCAATTTTGGCAAGGAGAGCTCGAGAAATATTGG 2280
2281 QY GACTGCTATATGGTAAACACATTTGCCCTGTGCTCTTCCACCACTGCATCTCTAACTGC 2340
2281 Db GATTGTCGATGGTGAAGCACAATTTGCTGTGCTCTTCCGCAACTGCATCTCTTACTGC 2340
2341 QY CCGTGGCTTCTCTGCT 2400
2341 Db CCGTGGCTTCTCTTATCT 2400
2401 QY ATTAAGTCTTATCT 2460
2401 Db ATTAAGTCTTATCT 2460
2461 QY ATCTGTTCAATCT 2520
2461 Db ATCTGTTCAATCT 2520
2521 QY TGGCAAGATCAAAACACCCAGCTTGAATTTCAATTTACTCTCTGATGATCTCGAAGACG 2580
2521 Db TGGATGAGATCAAAACACCCAGCTTGAATTTCAATTTACTCTCTGATGATCTCGAAGACG 2580
2581 QY TCCTGTGACTCAACTCAAGCCCTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACTCT 2640
2581 Db TCCTGTGACTCAACCCAGCCCTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACTCT 2640
2641 QY CTTCCAGTTCCTGCTTCACTCCAGCTTATCCAGTCTGATGAGTGCATCTCTCTCTCT 2700
2641 Db CTTCCAGTTCCTGCTTCACTCCAGCTTATCCAGTCTGATGAGTGCATCTCTCTCTCTCT 2700
2701 QY GTGGCATTGCTCCCATGCTCTTA 2723
2701 Db GTTCATTGCTCCCATGCTCTTA 2723

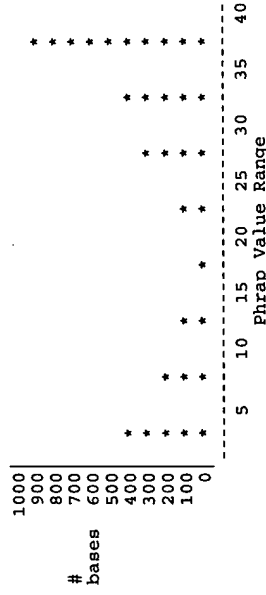
RESULT 13
AK075399
LOCUS
DEFINITION Homo sapiens cDNA PSEC0089 fis, clone PLACE100148, highly similar to Homo sapiens orphan G protein-coupled receptor HG38 mRNA.
ACCESSION AK075399
VERSION AK075399.1 GI:22761463
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Ota, T., Nishikawa, T., Suzuki, Y., Kawai-Hio, Y., Hayashi, K., Ishii, S., Saito, K., Yamamoto, J., Wakamatsu, A., Nagai, T., Nakamura, Y., Sato, H., Nagahara, K., Sugano, S. and Isogai, T.
TITLE HRI human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4095)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2002) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT HRI human cDNA sequencing project; cDNA 5'- & 3'-end one pass sequencing, clone selection and full insert sequencing; Helix Research Institute (supported by Japan Key Technology Center etc.); cDNA library construction; Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center.
FEATURES
source 1..4095
/organism="Homo sapiens"
/mol_type="mRNA"

Phrap values in estimate:
Average error rate (BCM-Phrap estimate):
Fraction of Phrap values less than 40 :
Number of consensus changing edits:
Number of N's in consensus :

144455
5.28339e-05
0.0137759
19
0

Position	Consensus changing edits	
	Original+Context	Edited+Context
6882	agatgggat(g)gcgtggcctt	agatgggat(g)gcgtggcctt
12828	tttcggtctc(n)tgcaactccc	tttcggtctc(c)tgcaactccc
13151	aaacacacat(n)caatagataa	aaacacacat(a)caatagataa
28957	tgtatattga(n)atggcattctt	tgtatattga(a)atggcattctt
37139	acgtgtatgt(n)tggttaatttt	acgtgtatgt(a)tggttaatttt
58030	ttttttattt(n)tatagttaatt	ttttttattt(t)tatagttaatt
58290	tcatgacctt(n)tcatgtccctt	tcatgacctt(a)tcatgtccctt
64729	ccatctctac(n)aaagaaaga	ccatctctac(g)aaagaaaga
92472	aatgggtgtt(g)aaatgggccc	aatgggtgtt(t)aaatgggccc
92662	cttagactgt(n)aatgcacgga	cttagactgt(t)aatgcacgga
103773	gctcaatttt(n)catttttagt	gctcaatttt(a)catttttagt
104056	ccatcatgct(n)ggtaatttgtc	ccatcatgct(a)ggtaatttgtc
117534	tcaagttgtg(n)atttcaggtt	tcaagttgtg(a)atttcaggtt
117581	cgaatcatct(n)agggaagctt	cgaatcatct(g)agggaagctt
125485	gcttcacctt(n)aaaatgaacc	gcttcacctt(g)aaaatgaacc
133723	gcatgggtgt(n)nncaactaat	gcatgggtgt(t)nncaactaat
133724	gcattgtgtt(n)ncacttaata	gcattgtgtt(g)ncacttaata
133725	catgtgtgtt(n)cacttaataa	catgtgtgtt(c)cacttaataa
133725	aaatatgttc(n)taataaagt	aaatatgttc(c)taataaagt
140801		

----- Distribution of Quality < 40 Bases



```

FEATURES
Source
Version: 1.01 qxf0.
Location/Qualifiers
1. .145165
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="12"
/clone="RP11-186F10"
1. .2004
/misc_feature
/note=".overlaps bases 52259..54262 of clone AC073612"
/function="clone overlap"
repeat_region
complement(6..223)
/rpt_family="LIME1"
repeat_region
complement(404..698)
/rpt_family="LIM4"
repeat_region
complement(833..1077)
/rpt_family="LIM4"
repeat_region
complement(1228..1482)
/rpt_family="LIM4"
repeat_region
complement(1740..2021)
/rpt_family="AluJb"
repeat_region
complement(2098..2151)
/rpt_family="LIMEc"
repeat_region
complement(2152..2449)

```


Qy	1334	TAAC	TCACTTA	AAAATTAA	CAGGA	AAATCAT	TGCGCTTAC	AGAGCTTG	TGATATCAT	TGTGA	AAACT	1339
Db	1334	TGAT	GCACCT	TGAAGCT	CA	AAGGGAACTTGGCCCTGTCT	CAAGGCCTTCT	CAAGGACACGATTT				1339
Qy	1394	TTCC	AGAACTCA	AGGTTATAG	AAATGCTTAT	GTCTATAC	CTTAC	AGTGTGCTG	CAATTTGGAGTGT			1453
Db	1394	TCCC	AAAAC	TGAGGATCCT	GGAGGTGC	CTTAC	GGCTTAC	AGTGTCTG	TGTCCTAC	TGCGCTAC	TGCGCATCT	1453
Qy	1454	GTG	AGAATG	CCCTATA	TAA	GATTTTCT	AAATCA	ATGGAATAA	AGGTGAC---	AA	CAGCAGTATGG	1510
Db	1454	GTGC	CAGCTTCTT	CA	AGACCTCTCGG	CAGTGG	CAGCGCCG	AGGACTTTT	CTCATCC	CA	AGAAG	1513
Qy	1511	ACG	ACCTTCA	TAGAA	AGATGCT	CGGAATGTTT	CA	GGCTCA	AGATGA	AGCTG	CACTTGAAG	1570
Db	1514	AGG	AGCACCA	AAAGAGG	CCCTGGGTCT	CTCTGT	CTG	GACCAAGCT	TGA	ACCATAT	TGACC	1573
Qy	1571	ATT	TCCTGCT	TGACTTT	TGAG-----	GA	AGACCTG	AAAGCCCTT	CA	TTCA	CTGTCAGT	1621
Db	1574	TAG	ACCTT	GGATG	AGCTCC	AGATGG	GGACAG	AGGACTCAA	AGCCAA	CCCC	CAGTGTCC	1633
Qy	1622	GTT	CACTTCC	CC	CAGGCCCCCTT	CA	AAACCTCTG	TGAACAC	CTCTG	ATGATGG	CTTGTGATCA	1681
Db	1634	GC	AGCCCTGT	CT	CAGGCCCCCTT	CA	AGCCCTG	CGAGAC	CTCTT	TGAGAGCT	TGGGCATCC	1693
Qy	1682	GA	ATTG	AGATGTG	GA	CAATAG	CAGTCTT	GGC	ACTTACTT	GTAA	TGCTTGGTGACTTCAA	1741
Db	1694	GC	TTTGCTGT	TGGGCC	CAATCGT	GTCTCT	CGTACTCTG	CTA	AGCGGCTG	GTCTGT	CTGTGA	1753
Qy	1742	CAG	TTTTT---	CAG	ATCCCTCTGT	TAC	ATTTTCC	CCCATTTAA	ACTGTTAA	TTTGGG	GTATCG	1798
Db	1754	CAG	CTTTT	TG	CAGCGG	CCCC	AGCCGCTGT	CTCC	CCGTC	CAAGCTT	TGGTGGGTGCGATGG	1813
Qy	1799	CAG	CAGTGA	CAATGCT	CA	CGGGAGTCT	C	CAGTGC	GTGCTGG	CTGTG	TGATGCGTGTCA	1858
Db	1814	CAG	GGCCAA	CGCCCTG	AC	GGGCAATTTCT	TGTGTCTCT	TG	GTCTCT	TG	TGGA	1873
Qy	1859	CTTT	TGGCAGCTTT	GC	ACG	ACATGGTGC	CTGTGG	GAGAA	TGGGGTT	TGGTT	GGCATGTCA	1918
Db	1874	CCT	ATGGT	CAGTT	TCGCT	GAGTATG	AGACCCGCT	GGGAG	AGCGGTCTGG	GTCTG	CCAGGCTA	1933
Qy	1919	TTG	TTTTTTT	TGTCC	ATTTTTT	TGCTT	CAGAA	TCACTG	TTTTTCT	TGCTT	ACTCTG	1978
Db	1934	CGG	CTTCT	TGGCTGT	CT	TGGTTT	CA	AGGCGT	CGGTGCT	CT	CACACTGG	1993
Qy	1979	TGG	AGCTGG	GGTTCCT	CTGT	GAAATAT	TCTG	CAAAATTTG	AAAGCA	AGCTTCC	ATTTTCTA	2038
Db	1994	TG	CAGTGC	AGCATCTCT	GT	GACCTG	CGT	CGAG	CTAC	GGGAGG	CGCGCTG	2053
Qy	2039	GC	CTGAA	AGTAA	TCA	TTTGT	CTGTG	CCCTG	CTG	GGCCTT	TGA	2098
Db	2054	CG	CTCCG	CGCAGG	CGC	ACTGG	GA	TGCTGG	CGCTG	CGCGG	CTG	2113
Qy	2099	TG	CTGGGT	GGCAG	CAAGTAT	TGG	CCCTCC	CTCTG	CTG	CTG	CTG	2152
Db	2114	TGG	CTCTGG	TGGGAGATG	TG	GGCGCT	CTCC	CACTCTG	CTG	CTG	CTG	2173
Qy	2153	GGG	AGCCAG	CACCA	TGGGCT	TACATG	TGCT	CGCTCAT	TTGCT	CAATTC	CAATTC	2212
Db	2174	GCC	GGCCG	GGCCG	CTG	GGCTT	CGCTG	TAG	CCCTG	GGTGA	TGA	2233
Qy	2213	TC	ATGATG	ACCATTT	GCCTT	AC	CAAGCTCT	ACTG	CAATTT	TGGA	CAAGG	2272
Db	2234	TGG	TGGTGG	CCGGG	CGCTTAC	ATCA	AGCTCT	ACT	TGTAG	CCCTGCC	ACG	2293
Qy	2273	ATA	TTTGG	GA	CTGTCTAT	TG	TAAACAC	ATTTGG	CCCTG	TTCTT	CA	2332
Db	2294	CCG	TGTGG	GA	CTGG	CGCATG	TGTG	CGC	CA	CGTGG	CTCAT	2353
Qy	2333	TAA	ACTGCC	CTG	TGGCTTCT	TGCTT	CTCT	CTCTT	TAA	TAA	CACTTAT	2392
Db	2354	TCT	ACTG	CCCCG	TGG	CCCTTCT	CT	CAGCTTT	TG	CCCTCA	TG	2413
Qy	2393	CT	GAAGTAA	TTTAA	GTTTAT	CTCT	CTG	TGGT	TAGTCC	CACTTCT	CTG	2452

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 12, 2005, 08:18:12 ; Search time 1573.87 Seconds
(without alignments)
10245.706 Million cell updates/sec

Title: US-10-751-736-21

Perfect score: 2724

Sequence: 1 atggacactccggctcgg.....cattgtcccatgtctctaa 2724

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2724	100.0	2724	3	AAA30770 Human G p
2	2724	100.0	2724	8	ABZ42816 Human G p
3	2724	100.0	2724	10	ADC22782 Human G p
4	2724	100.0	2724	10	ADH14255 Human HG3
5	2724	100.0	2724	12	ADI32985 Human G p
6	2724	100.0	2724	12	ADO29923 Human GPC
7	2724	100.0	2880	11	ADN33796 Cancer/an
8	2724	100.0	2880	13	ADQ80249 G protein
9	2724	100.0	2373	13	ADR67869 Human HG3
10	2724	100.0	3032	11	ADN39627 Cancer/an
11	2724	100.0	3032	11	ADN39530 Cancer/an
12	2724	100.0	4558	2	AAx233980 Human HG3
13	2724	100.0	4570	12	ADL12472 Human ste
14	2720.8	99.9	2724	3	AAA30779 DNA encod
15	2720.8	99.9	2724	10	ADC22796 Human G p
16	2720.8	99.9	2724	10	ADH14269 Mutated h
17	2719.4	99.8	3438	10	ADF70582 Orphan re
18	2712	99.6	3297	2	AAx233981 Human HG3
19	2566	94.2	2651	10	ADB80463 Ovarian c
20	2566	94.2	2651	11	ADN39165 Cancer/an

21	2566	94.2	2651	11	ADN39795	Adn39795 Cancer/an
22	2074.6	76.2	2082	2	AAZ25344	Aaz25344 Human LGR
23	2052.6	75.4	2724	12	ADO30213	Ado30213 Mouse GPC
24	2052.6	75.4	3098	13	ADR67870	Adr67870 Mouse HG3
25	996	36.6	1790	6	ABQ55077	Abq55077 Human ova
26	841.4	30.9	2901	6	AAI67921	Aai67921 Mouse LGR
27	841.4	30.9	2901	11	ADN02241	Adn02241 Mouse CDS
28	841.4	30.9	2901	12	ADK19406	Adk19406 ORF of CD
29	841.4	30.9	3637	6	AAI67920	Aai67920 Mouse LGR
30	841.4	30.9	3637	11	ADN02239	Adn02239 Mouse CDN
31	841.4	30.9	3637	12	ADK19404	Adk19404 cDNA enco
32	813.2	29.9	2901	6	AAI67927	Aai67927 Human LGR
33	813.2	29.9	2901	10	ADC16694	Adc16694 cDNA enco
34	813.2	29.9	2901	10	ADC16692	Adc16692 cDNA enco
35	813.2	29.9	2901	10	ADC16696	Adc16696 cDNA enco
36	813.2	29.9	2901	11	ADN02250	Adn02250 Human ful
37	813.2	29.9	2901	12	ADK19415	Adk19415 ORF of fu
38	813.2	29.9	3438	10	ABT31938	Abt31938 Human bre
39	813.2	29.9	3443	10	ADD89090	Add89090 Encoding
40	813.2	29.9	3492	6	AAI67926	Aai67926 Human LGR
41	813.2	29.9	3492	11	ADN02248	Adn02248 Human ful
42	813.2	29.9	3492	12	ADK19413	Adk19413 Full leng
43	813.2	29.9	3492	12	ADQ15073	Adq15073 Human can
44	813.2	29.9	3618	10	ADF70543	Adf70543 Orphan re
45	811.6	29.8	2901	10	ADC16695	Adc16695 cDNA with

ALIGNMENTS

RESULT 1
AAA30770
ID AAA30770 standard; cDNA; 2724 BP.
XX
AC AAA30770;
XX
DT 21-AUG-2000 (first entry)
XX
DE Human G protein-coupled receptor HG38 cDNA.
XX
KW G protein-coupled receptor; GPCR; constitutively active;
KW intracellular loop 3; transmembrane domain 6; drug screening; agonist;
KW antagonist; ss.
XX
OS Homo sapiens.
XX
PN WO200022129-A1.
XX
PD 20-APR-2000.
XX
PF 12-OCT-1999; 99WO-US023938.
XX
PR 13-OCT-1998; 98US-00170496.
XX
(AREN-) ARENA PHARM INC.
XX
Behan DP, Chalmers DT, Liaw CW;
XX
WPI: 2000-329165/28.
XX
P-PSDB; AA90682.
XX
Non-endogenous constitutively activated human G protein-coupled
PT receptors, useful for identifying agonists for use as pharmaceutical
agents.
XX
Example 1; Page 315-317; 341pp; English.
XX

The invention relates to constitutively active, non-endogenous versions of endogenous human orphan G protein-coupled receptors (GPCRs, AA90643- AA90677 and AA90683-Y90687) and to DNA encoding them (AAA30709-AA30743 and AAA30775-AA30779). The mutant proteins of the invention contain a mutation in a portion of the protein comprising intracellular loop 3 (IC3) and transmembrane domain 6 (TM6). A non-endogenous amino acid, X,

CC is substituted for an endogenous residue in IC3 at a position 16 amino
CC acids N-terminal of an endogenous proline in TM6 to form a sequence X-
CC (AA)15-Pro. The endogenous amino acid is selected from Lys, His, Arg or
CC Ala, and is preferably Lys. When the endogenous residue at this position
CC is Lys, this residue is replaced by His, Arg or preferably Ala. The 15
CC amino acid stretch between the substituted amino acid and the Pro may be
CC endogenous, non-endogenous, or a mixture of endogenous and non-endogenous
CC residues. The constitutively active GPCRs are useful for identifying
CC antagonists, agonists and partial agonists for use as pharmaceutical
CC agents. The mutant proteins are also useful in research settings for
CC elucidating the roles of the receptors in normal and diseased conditions.
CC Antagonists for a particular GPCR are useful for treating diseases and
CC disorders associated with that receptor. Because the novel mutant GPCRs
CC are constitutively active, they can be used directly for screening of
CC compounds without the need for endogenous ligands. The present sequence
CC represents cDNA encoding a human wild-type GPCR used in an
CC exemplification of the invention. This was cloned and subjected to site-
CC directed mutagenesis (SDM) to generate DNA encoding the corresponding
XX mutant of the invention

SQ Sequence 2724 BP; 665 A; 723 C; 554 G; 782 T; 0 U; 0 Other;

Query Match 100.0%; Score 2724; DB 3; Length 2724;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGACACTCCGGCTCGGTGTCTCTGTCTCTTGGCTGTGCTGCTGAGCTGGCGACC 60
DB 1 ATGGACACTCCGGCTCGGTGTCTCTGTCTCTTGGCTGTGCTGCTGAGCTGGCGACC 60
QY 61 GGGGCGAGCTCTCCAGGCTGGTGTGTTGCTGAGGGGCTGCCACACACTGTCATTGC 120
DB 61 GGGGCGAGCTCTCCAGGCTGGTGTGTTGCTGAGGGGCTGCCACACACTGTCATTGC 120
QY 121 GAGCCCGAGCGGAGGATTTGCTCAGGGTGGACTGCTCCGACTGGGGCTCTCGGAGCTG 180
DB 121 GAGCCCGAGCGGAGGATTTGCTCAGGGTGGACTGCTCCGACTGGGGCTCTCGGAGCTG 180
QY 181 CTTTCCAACTCAGGCTCTACCTCTTACCTAGACCTCAGTATGAAACAATCAGTCAG 240
DB 181 CTTTCCAACTCAGGCTCTACCTCTTACCTAGACCTCAGTATGAAACAATCAGTCAG 240
QY 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTTCTGAGGAGGTACGCTTTGCGGGAAC 300
DB 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTTCTGAGGAGGTACGCTTTGCGGGAAC 300
QY 301 GCTCTGACATACATTTCCCAAGGGAGCAATTCATGGCCCTTTACAGTCTTTAAAGTCTTATG 360
DB 301 GCTCTGACATACATTTCCCAAGGGAGCAATTCATGGCCCTTTACAGTCTTTAAAGTCTTATG 360
QY 361 CTGCGAATATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGGGAGCCTT 420
DB 361 CTGCGAATATCAGCTAAGACACGTACCCACAGAGCTCTGCAGAAATTTGGGAGCCTT 420
QY 421 CAATCCCTCGCTCGGATGCTAAACACATCAGTATGCCCCCAAGCTGTTTCAGTGCG 480
DB 421 CAATCCCTCGCTCGGATGCTAAACACATCAGTATGCCCCCAAGCTGTTTCAGTGCG 480
QY 481 CTGCATTTCCCTGAGGCACCTGTGGCTGGATGCAATGGGTTAAACAGAAATCCCGTCCAG 540
DB 481 CTGCATTTCCCTGAGGCACCTGTGGCTGGATGCAATGGGTTAAACAGAAATCCCGTCCAG 540
QY 541 GCTTTTAGAAGTTATCGGCATTGCAAGCATGACCTTGGCCCTGAAACAATAACACCAC 600
DB 541 GCTTTTAGAAGTTATCGGCATTGCAAGCATGACCTTGGCCCTGAAACAATAACACCAC 600
QY 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGTAGTCTACATCTCCATAACCAAT 660
DB 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGTAGTCTACATCTCCATAACCAAT 660
QY 661 AGAATCCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
DB 661 AGAATCCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720

QY 721 TTAAATTACATAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
DB 721 TTAAATTACATAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
QY 781 GAACTAGGATTTCAAGCAACAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB 781 GAACTAGGATTTCAAGCAACAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
QY 841 CCTTCTCTTATTAACAATCAATTTCTATGACAATCCCATCCAATTTGTTGGGAGATCTGCT 900
DB 841 CCTTCTCTTATTAACAATCAATTTCTATGACAATCCCATCCAATTTGTTGGGAGATCTGCT 900
QY 901 TTTCAACATTTACTCTGAACTAAGAACACTGAACTGAAATGGTGGCTTCAAAATAACTGAA 960
DB 901 TTTCAACATTTACTCTGAACTAAGAACACTGAACTGAAATGGTGGCTTCAAAATAACTGAA 960
QY 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
DB 961 TTTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
QY 1021 TCACTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
DB 1021 TCACTCTTCTCAAAACCGTCTGCAATCAGTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
QY 1081 TACAACCTATTAGAGATTTTACCAGTCTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1140
DB 1081 TACAACCTATTAGAGATTTTACCAGTCTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1140
QY 1141 CTAAGACATAATGAAATCTAGCAAAATTAAGTGTGACACTTTCCAGCAGTGTGCTTAGCCTC 1200
DB 1141 CTAAGACATAATGAAATCTAGCAAAATTAAGTGTGACACTTTCCAGCAGTGTGCTTAGCCTC 1200
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTCCACCCCAATGSCATTTTCCACT 1260
DB 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATTTGCTATTATTCCACCCCAATGSCATTTTCCACT 1260
QY 1261 TTGCGATCCCTTAATAAGCTGGACCTATCGTCCCAACCTCTGCTGCTTTTCCCTATAACT 1320
DB 1261 TTGCGATCCCTTAATAAGCTGGACCTATCGTCCCAACCTCTGCTGCTTTTCCCTATAACT 1320
QY 1321 GGGTTACATGCTGTTAACTCACTTAAATTAACAGAAATCAATGCTTACAGAGCTTGATA 1380
DB 1321 GGGTTACATGCTGTTAACTCACTTAAATTAACAGAAATCAATGCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAATCTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGAGCTGTG 1440
DB 1381 TCATCTGAAATCTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGAGCTGTG 1440
QY 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGCAACAC 1500
DB 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGCAACAC 1500
QY 1501 AGCAGTATGGAGCAGCTTTTCAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACT 1560
DB 1501 AGCAGTATGGAGCAGCTTTTCAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACT 1560
QY 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAAGCTTCAAGAGCCCTTCAATTCAGTGAG 1620
DB 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAAGCTTCAAGAGCCCTTCAATTCAGTGAG 1620
QY 1621 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGCTGATGGCTGGCTGATC 1680
DB 1621 TGTTCACCTTCCCAAGGCCCTTCAAAACCTGTGAAACACCTGCTGATGGCTGGCTGATC 1680
QY 1681 AGAATGGAGTGTGGACCATAGCAGTCTGCGACCTTACTTGTAACTGCTTGTGATCTTCA 1740
DB 1681 AGAATGGAGTGTGGACCATAGCAGTCTGCGACCTTACTTGTAACTGCTTGTGATCTTCA 1740
QY 1741 ACAGTGTTCAGATCCCTCTGTACATTTCCCGCCATTAACCTGTTAACTGTTGGGGTCAACGA 1800
DB 1741 ACAGTGTTCAGATCCCTCTGTACATTTCCCGCCATTAACCTGTTAACTGTTGGGGTCAACGA 1800

Qy		1801	GCAGTGAACATGCTCA CGGGAGTCTCAGATGCCGTGCTGGTGTTGGATGCGTTCACT	1861
Db		1801	GCAGTGAACATGCTCA CGGGAGTCTCAGTATGCCGTGCTGGTGTTGGATGCGTTCACT	1860
Qy		1861	TTTGGCAGCTTTGCACGACATGGTCCTGGTGGGAGAAATGGGGTTGGTGCCTCATGTCATT	1920
Db		1861	TTTGGCAGCTTTGCACGACATGGTCCTGGTGGGAGAAATGGGGTTGGTGCCTCATGTCATT	1920
Qy		1921	GGTTTTTTTGCCATTTTTTGCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGAGCCCTG	1980
Db		1921	GGTTTTTTTGCCATTTTTTGCTTCAGAAATCATCTGTTTTCTGCTTACTCTGGCAGAGCCCTG	1980
Qy		1981	GAGCGTGGGTTCTCTGTGAAATAATTTGCGAAAATTTGAAAACGAAAGCTCCAATTTCTTAGC	2040
Db		1981	GAGCGTGGGTTCTCTGTGAAATAATTTGCGAAAATTTGAAAACGAAAGCTCCAATTTCTTAGC	2040
Qy		2041	CTGAAAGTAATCATTTTTGCTCTGTGCCCTCTGCTGGCCCTTGACCATGGCCGAGATTCGCCCTG	2100
Db		2041	CTGAAAGTAATCATTTTTGCTCTGTGCCCTCTGCTGGCCCTTGACCATGGCCGAGATTCGCCCTG	2100
Qy		2101	CTGGGTGGCAGCAGTAGTAGGCGCCTCCCCTCTCTGCTGCTTGGCTTTGGCGGAGCCCC	2160
Db		2101	CTGGGTGGCAGCAGTAGTAGGCGCCTCCCCTCTCTGCTGCTTGGCTTTGGCGGAGCCCC	2160
Qy		2161	AGCACATAGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG	2220
Db		2161	AGCACATAGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG	2220
Qy		2221	ACCAATTGCCCTACACCAAGCTCTA CTGCAATTTGGACAAGGAGACCTCGGAGAATATTTGG	2280
Db		2221	ACCAATTGCCCTACACCAAGCTCTA CTGCAATTTGGACAAGGAGACCTCGGAGAATATTTGG	2280
Qy		2281	GACTGCTCTATGTTAAACACATTTGCCCTGTGTCTTTCACCAAAGCTCATCTTAAACTGC	2340
Db		2281	GACTGCTCTATGTTAAACACATTTGCCCTGTGTCTTTCACCAAAGCTCATCTTAAACTGC	2340
Qy		2341	CCTGTGCTCTTCTGTGCTTCTCTCTTTTAAATAAA CCTTACATTAATFACGTCCTGAAGTA	2400
Db		2341	CCTGTGCTCTTCTGTGCTTCTCTCTTTTAAATAAA CCTTACATTAATFACGTCCTGAAGTA	2400
Qy		2401	ATTAAAGTTTATCCTTCTGGTGGTAGTCCCACTTCCCTGCTCATGTCATCCCTTCTCTAC	2460
Db		2401	ATTAAAGTTTATCCTTCTGGTGGTAGTCCCACTTCCCTGCTCATGTCATCCCTTCTCTAC	2460
Qy		2461	ATCTTGTFTCAATCCTCACTTTTAAAGGAGGATCTGCTGAGCCCTGAGAAAGCAAA CTTACGTC	2520
Db		2461	ATCTTGTFTCAATCCTCACTTTTAAAGGAGGATCTGCTGAGCCCTGAGAAAGCAAA CTTACGTC	2520
Qy		2521	TGGACAAGATCAAAAACA CCCAAGCTTGATGTCAATTAAC TCTGATGATGCGAAAAACAG	2580
Db		2521	TGGACAAGATCAAAAACA CCCAAGCTTGATGTCAATTAAC TCTGATGATGCGAAAAACAG	2580
Qy		2581	TCCGTGACTCAACTCAAGCCTTGGTAA CTTTACCAGCTCCAGACATCACTTATGACCTG	2640
Db		2581	TCCGTGACTCAACTCAAGCCTTGGTAA CTTTACCAGCTCCAGACATCACTTATGACCTG	2640
Qy		2641	CCTCCCAAGTTCGGTGCCATCA CCAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCT	2700
Db		2641	CCTCCCAAGTTCGGTGCCATCA CCAGCTTATCCAGTGACTGAGAGCTGCCATCTTTCTCT	2700
Qy		2701	GTGGCATTTGTGCCATGTCTCTAA	2724
Db		2701	GTGGCATTTGTGCCATGTCTCTAA	2724

RESULT 2

RESUL 2
ABZ42816

ABZ42816
ID ABZ42816 standard: DNA: 2724 BP.

XX
01074784

AC ABZ42816.

AC
XX
ABZ42816;

XX
01-MAR-2003 (First entry)DT 04-MAR-2003 (first entry)
yy[illegible]

Human G protein-coupled receptor GPR49 nucleotide SEQ ID NO:421.

G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
G protein-coupled receptor modulator; antibody; immune-related disease;
growth-related disease; cell regeneration-related disease; AIDS; cancer;
immunological-related cell proliferative disease; autoimmune disease;
Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
ulcer; gene; ds.

Homo sapiens.

WO200261087-A2.

08-AUG-2002.

19-DEC-2001; 2001WO-US050107.

19-DEC-2000; 2000US-0257144P.

(LIFE-) LIFESPAN BIOSCIENCES INC.

Burmer GC, Roush CL, Brown JP;
WPI, 2003-046718/04.
P-PSDB; ABP1969.

New isolated antigenic peptides e.g., for G protein-coupled receptors (GPCR), useful for diagnosing and designing drugs for treating conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or autoimmune diseases.

Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptides (I) comprising: (a) any one of 1601 sequences (see ABP2019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (I) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease, atherosclerosis, bacteremia, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. ABP242523 to ABP242869 encode GPCR proteins given in ABP16175 to ABP82018, which are used in the exemplification of the present invention.

1

1

1

1

ف

Qy

১৫

Dib.

•

QY 61 GGGGGCAGCTCTCCAGAGTCTGGTGTGTGCTGAGGGGCTGCCACACACACTGTCTATGTC 120
Db |||||
QY 61 GGGGGCAGCTCTCCAGAGTCTGGTGTGTGCTGAGGGGCTGCCACACACACTGTCTATGTC 120
Db |||||
QY 121 GAGCCCGAGCGGACGAGATGTTGCTCAGGGTGGAGTCTCCGACCTGGGGCTCTCGGAGCTG 180
Db |||||
QY 121 GAGCCCGAGCGGACGAGATGTTGCTCAGGGTGGAGTCTCCGACCTGGGGCTCTCGGAGCTG 180
Db |||||
QY 181 CTTTCCAACTCAGCGTCTTCACTCTCCTACCTAGACCTCAGTATGAACAAACATCAGTCA 240
Db |||||
QY 241 CTGCTCCCGAATCCCTCCGACGATCTCCGCTTCCTGGAGGAGTACGTTCCGGGAAAC 300
Db |||||
QY 241 CTGCTCCCGAATCCCTCCGACGATCTCCGCTTCCTGGAGGAGTACGTTCCGGGAAAC 300
Db |||||
QY 301 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCCTTTACAGTCTTAAAGTCTCTATG 360
Db |||||
QY 301 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCCTTTACAGTCTTAAAGTCTCTATG 360
Db |||||
QY 361 CTGCAGAAATAATCAGCTAAGACACGTACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db |||||
QY 421 CAATCCCTGCGTCTGGATGCTTAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db |||||
QY 481 CTGCATTTCCCTGAGGCACTGTGGCTGATGACATGCGTTAAACAGAAATCCCGTCCAG 540
Db |||||
QY 541 GCTTTTGAAGTTTATCGGCAATGCAAGCCATGACCTTGGCCCTGAAACAAATAACCCAC 600
Db |||||
QY 601 ATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCGATACAAAT 660
Db |||||
QY 601 ATACAGACTATGCCCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCGATACAAAT 660
Db |||||
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Db |||||
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Db |||||
QY 721 TTAATATCAATAAATCCTTTGATGAAATTCGCCACTGCAATTTAGGACACTCTCCAAACCTTAA 780
Db |||||
QY 781 GAACTAGGATTTATAGCAACAAATACAGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db |||||
QY 781 GAACTAGGATTTATAGCAACAAATACAGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db |||||
QY 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Db |||||
QY 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Db |||||
QY 901 TTTCAACATTTTACCTGAACTAAGAACACTGACTCTGAATGGTGCCTCAAAATAACTGAA 960
Db |||||
QY 961 TTTCCCTGATTTAACTGGAACTGCAACCTGGAGACTGCACTTTAACTGGAGCACAGATC 1020
Db |||||
QY 961 TTTCCCTGATTTAACTGGAACTGCAACCTGGAGACTGCACTTTAACTGGAGCACAGATC 1020
Db |||||
QY 1021 TCATCTCTTCCCTCAAAACGCTCTGCAATCAGTTTACCTAACTCTCAAGTGTCTAGATCTGCT 1080
Db |||||
QY 1021 TCATCTCTTCCCTCAAAACGCTCTGCAATCAGTTTACCTAACTCTCAAGTGTCTAGATCTGCT 1080
Db |||||
QY 1081 TACAACCTTATTAGAAGATTTACCCAGTCTGCAATTTTCAAGTGTGCAAAAGCTTCAGAAATTCAC 1140
Db |||||
QY 1081 TACAACCTTATTAGAAGATTTACCCAGTCTGCAATTTTCAAGTGTGCAAAAGCTTCAGAAATTCAC 1140
Db |||||

QY 1141 CTAAGACATAATGAAATCTACGAAATTAAGTGTGACACTTTTCCAGCAGTGTCTTAGCCTC 1200
Db |||||
QY 1141 CTAAGACATAATGAAATCTACGAAATTAAGTGTGACACTTTTCCAGCAGTGTCTTAGCCTC 1200
Db |||||
QY 1201 CGATCGCTGAAATTTGGCTTGGAAACAAATTTCTATTATTCACCCCAATTCGCAATTTTCCACT 1260
Db |||||
QY 1201 CGATCGCTGAAATTTGGCTTGGAAACAAATTTCTATTATTTACCCCAATTCGCAATTTTCCACT 1260
Db |||||
QY 1261 TTGCCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTCTGTCTTTTCCCTATAACT 1320
Db |||||
QY 1261 TTGCCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTCTGTCTTTTCCCTATAACT 1320
Db |||||
QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGTATA 1380
Db |||||
QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGAAATCATGCTTACAGAGCTTGTATA 1380
Db |||||
QY 1381 TCATCTGAAATCTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTCTGT 1440
Db |||||
QY 1381 TCATCTGAAATCTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTCTGT 1440
Db |||||
QY 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
Db |||||
QY 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
Db |||||
QY 1501 AGCAGTATGGAGACCTTTTCAAGAAAGATGCTCGAAATGTTTTCAGGCTCAAGATGAACTG 1560
Db |||||
QY 1501 AGCAGTATGGAGACCTTTTCAAGAAAGATGCTCGAAATGTTTTCAGGCTCAAGATGAACTG 1560
Db |||||
QY 1561 GACCTTGAAGATTTTCTGCTTGTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db |||||
QY 1561 GACCTTGAAGATTTTCTGCTTGTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db |||||
QY 1621 TGTTCACCTTCCCAAGGCCCTTCAAACCTGTGAAACACTGCTGTGATGGCTGGCTGATC 1680
Db |||||
QY 1621 TGTTCACCTTCCCAAGGCCCTTCAAACCTGTGAAACACTGCTGTGATGGCTGGCTGATC 1680
Db |||||
QY 1681 AGAATGGAGTGTGGACCATAGCAGTCTGCGCACTTACTGTAAATGCTTTGGTGACTTCA 1740
Db |||||
QY 1681 AGAATGGAGTGTGGACCATAGCAGTCTGCGCACTTACTGTAAATGCTTTGGTGACTTCA 1740
Db |||||
QY 1741 ACAGTTTTTACAGATCCCTCTGTACATTTTCCCAATTTAACTGTTAAATTTGGGGTCAATCGCA 1800
Db |||||
QY 1741 ACAGTTTTTACAGATCCCTCTGTACATTTTCCCAATTTAACTGTTAAATTTGGGGTCAATCGCA 1800
Db |||||
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCGTGTGGAACACTGCTGTGATGGCTGATC 1860
Db |||||
QY 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCGTGTGGAACACTGCTGTGATGGCTGATC 1860
Db |||||
QY 1861 TTTGGCAGCTTTGCAAGCATGCTGCTGGGAGAAATGGGGTTGGTGGCATGTCAAT 1920
Db |||||
QY 1861 TTTGGCAGCTTTGCAAGCATGCTGCTGGGAGAAATGGGGTTGGTGGCATGTCAAT 1920
Db |||||
QY 1921 GGTTTTTTGTCCATTTTTTGTTCAGAAATCACTGTCTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db |||||
QY 1921 GGTTTTTTGTCCATTTTTTGTTCAGAAATCACTGTCTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db |||||
QY 1981 GAGCGTGGGTTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTTCCATTTTCTAGC 2040
Db |||||
QY 1981 GAGCGTGGGTTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTTCCATTTTCTAGC 2040
Db |||||
QY 2041 CTGAAAGTAAATCAATTTTGTCTCTGTGCCCTTGACCATGCGCCGACAGTTCCTCCCTG 2100
Db |||||
QY 2041 CTGAAAGTAAATCAATTTTGTCTCTGTGCCCTTGACCATGCGCCGACAGTTCCTCCCTG 2100
Db |||||
QY 2101 CTGGTGGCAGCAAGATGCGCCCTCTCTGCTCTGCTTTGCGCTTTTGGGGAGGCC 2160
Db |||||
QY 2101 CTGGTGGCAGCAAGATGCGCCCTCTCTGCTCTGCTTTGCGCTTTTGGGGAGGCC 2160
Db |||||
QY 2161 AGCACCATTGGCTACATGCTCGCTCATCTGCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db |||||
QY 2161 AGCACCATTGGCTACATGCTCGCTCATCTGCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db |||||
QY 2221 ACCATTGCTTACACCAAGCTCTACTGCAATTTTGGACAAAGGGAGACCTGGAGAAATTTTGG 2280
Db |||||

Db 2221 ACCATGGCTACCAAGCTCTACTGCAATTTGGCAAGGGAGACCTGGAGAAATATTGG 2280
2281 GACTGCTCTATGTTAAACACATTTGCCCTTGCTCTTCCACCAACTGCATCTAAACTGC 2340
2281 GACTGCTCTATGTTAAACACATTTGCCCTTGCTCTTCCACCAACTGCATCTAAACTGC 2340
2341 CCGTGGCTTTCTGCT 2400
2341 CCGTGGCTTTCTGCT 2400
2401 ATTAAGTTTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
2401 ATTAAGTTTATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2460
2461 ATCTGTTTCAATCTCACTTTAAGGAGGATCTGTGAGCTGAGAAAGCAAACTTACGTC 2520
2461 ATCTGTTTCAATCTCACTTTAAGGAGGATCTGTGAGCTGAGAAAGCAAACTTACGTC 2520
2521 TGGCAAGATCAAAACACCAAGCTTGATGTCATTAATTAATTAATTAATTAATTAATTA 2580
2521 TGGCAAGATCAAAACACCAAGCTTGATGTCATTAATTAATTAATTAATTAATTAATTA 2580
2581 TCCTGTGACTCAACTCAAGCCTTGCTTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2640
2581 TCCTGTGACTCAACTCAAGCCTTGCTTAACCTTTACAGCTCCAGCATCACTTATGACCTG 2640
2641 CTTCCAGTTCGCTGCCATCACCAGCTTATCCAGTCACTGAGAGCTGCCATCTTTCCTCT 2700
2641 CTTCCAGTTCGCTGCCATCACCAGCTTATCCAGTCACTGAGAGCTGCCATCTTTCCTCT 2700
2701 GTGGCAATTTGCTCCATGCTCTTAA 2724
2701 GTGGCAATTTGCTCCATGCTCTTAA 2724

RESULT 3
ADC22782
ID ADC22782 standard; cDNA; 2724 BP.
XX
AC ADC22782;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human G protein-coupled receptor cDNA #39.
XX
KW Human; gene; ss; G protein-coupled receptor; GPCR;
KW transmembrane-6 region; TM6; intracellular-3 region; IC3.
XX
OS Homo sapiens.
XX
PN US655339-B1.
XX
PD 29-APR-2003.
XX
PF 13-OCT-1998; 98US-00170496.
XX
PR 14-APR-1997; 97US-00839449.
PR 14-APR-1998; 98US-00060188.
PR 26-JUN-1998; 98US-0090783P.
PR 07-AUG-1998; 98US-0095677P.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Liaw CW, Behan DP, Chalmers DT;
XX
XX WPI; 2003-742861/70.
DR P-PSDB; ADC22783.
XX
PT Creating a constitutively active version of an endogenous human G protein
PT coupled receptor (GPCR) comprises substituting a specific amino acid in
PT the transmembrane-6 region with a different amino acid, and testing for
PT constitutive activity.

XX Example 1; SEQ ID NO 263; 221pp; English.
XX
CC The invention relates to a method for treating a non-endogenous,
CC constitutively active version of an endogenous human G protein-coupled
CC receptor (GPCR) that has a transmembrane-6 (TM6) region and an
CC intracellular-3 (IC3) region, by substituting a specific amino acid in
CC the TM6 region with a different amino acid, and testing for constitutive
CC activity. The method is useful for creating a constitutively active
CC version of an endogenous human GPCR that comprises a transmembrane 6
CC region and an intracellular loop 3 region. The altered human GPCR
CC polypeptides are useful for screening test compounds for identification
CC of inverse agonists or partial agonists of GPCR polypeptides, which may
CC be used in biological research. A nucleic acid encoding the altered GPCR
CC may be used to create a transgenic animal expressing the altered GPCR.
CC The method allows screening for compounds that modulate the activity of a
CC human G protein-coupled receptor without the need for provision of a
CC ligand for the receptor. This is particularly useful in allowing
CC screening of compounds against orphan receptors for which no ligand is
CC currently known. This sequence represents cDNA encoding a human GPCR
CC polypeptide of the invention.
XX
SQ Sequence 2724 BP; 665 A; 723 C; 554 G; 782 T; 0 U; 0 Other;
Query Match 100.0%; Score 2724; DB 10; Length 2724;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGACACTCCCGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
DB 1 ATGGACACTCCCGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 61 GGGGGCAGCTCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCACACACTGTCTATTGC 120
DB 61 GGGGGCAGCTCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCACACACTGTCTATTGC 120
QY 121 GAGCCGACGTCCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCACACACTGTCTATTGC 180
DB 121 GAGCCGACGTCCTCCAGGCTCTGGTGTGCTGAGGGGCTGCCACACACTGTCTATTGC 180
QY 181 CTTTCCAACTCAGGCTCTTCACTTCCCTTACCTAGACCTCAGTATGAACAACTAGTCAG 240
DB 181 CTTTCCAACTCAGGCTCTTCACTTCCCTTACCTAGACCTCAGTATGAACAACTAGTCAG 240
QY 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCTGAGGAGGTACGTTCTCGGGAAC 300
DB 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTCTGAGGAGGTACGTTCTCGGGAAC 300
QY 301 GCTCTGACATACATTCCCAAGGGAGCATTCATGCGCTTTTACAGTCTTAAAGTTCTTATG 360
DB 301 GCTCTGACATACATTCCCAAGGGAGCATTCATGCGCTTTTACAGTCTTAAAGTTCTTATG 360
QY 361 CTGCAGATAATCAGCTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGGAGCCCTT 420
DB 361 CTGCAGATAATCAGCTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGGAGCCCTT 420
QY 421 CAATCCCTGCTCTGGATGCTTAAACACATCAGTATGTGCCCCCAAGCTGTTTCACTGAG 480
DB 421 CAATCCCTGCTCTGGATGCTTAAACACATCAGTATGTGCCCCCAAGCTGTTTCACTGAG 480
QY 481 CTGCATTTCCCTGAGGACCTCTGTGGCTGGATGACAAATGCGTTTAAAGAAATCCCGTCCAG 540
DB 481 CTGCATTTCCCTGAGGACCTCTGTGGCTGGATGACAAATGCGTTTAAAGAAATCCCGTCCAG 540
QY 541 GCTTTTAGAAGTTTATCGGCATTCGAGCCATGACCTTTGGCCCTGAAACAATATACCCAC 600
DB 541 GCTTTTAGAAGTTTATCGGCATTCGAGCCATGACCTTTGGCCCTGAAACAATATACCCAC 600
QY 601 ATACCAAGTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAAT 660
DB 601 ATACCAAGTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAAT 660

	Qy	2521	TGACAGATCAAAACACCCAGCTTGATGTCAATTAACTCTCATCATGTCGAAAAACAG	2580
	Dd	2521	TGACAGATCAAAACACCCAGCTTGATGTCAATTAACTCTCATCATGTCGAAAAACAG	2580
	Qy	2581	TCCTGTGACTCAAACCTCAAGCCTTGGTAACCTTTACCAGCTGCCAGCATCACTTATGACCTG	2640
	Dd	2581	TCCTGTGACTCAAACCTCAAGCCTTGGTAACCTTTACCAGCTGCCAGCATCACTTATGACCTG	2640
	Qy	2641	CCTCCCAGTTCCGTGCCATCACCGACTTATCCAAGTGAAGAGCTGCCATCTTTCTCT	2700
	Dd	2641	CCTCCCAGTTCCGTGCCATCACCGACTTATCCAAGTGAAGAGCTGCCATCTTTCTCT	2700
	Qy	2701	GTGGCATTTGTCCCATGCTCTTAA	2724
	Dd	2701	GTGGCATTTGTCCCATGCTCTTAA	2724
 RESULT 5 ADI32985				
ID	AD132985	standard; DNA; 2724 BP.		
XX	XX	AC	AC	AD132985;
XX	XX	22-APR-2004 (first entry)		
XX	XX	Human G protein-coupled receptor (GPCR) 49 DNA.		
DE	XX	G protein-coupled receptor 49; GPCR; neuroprotective; neural;		
KW	KW	endocrine system disorder; gene therapy; antisense therapy; human; ds.		
OS	XX	Homo sapiens.		
XX	XX	US2003235910-A1.		
PN	XX	25-DEC-2003.		
PD	XX	17-JUN-2002; 2002US-00174456.		
PP	XX	17-JUN-2002; 2002US-00174456.		
PR	XX	(ISIS-) ISIS PHARM INC.		
PA	XX	Monia BP, Freier SM;		
PI	XX	WPI; 2004-070584/07.		
DR	XX			
PT	PT	New antisense oligonucleotide comprising a sequence targeted to a nucleic acid encoding G protein-coupled receptor 49, useful for preparing a composition for treating e.g., neural or endocrine system disorder.		
PT	PT			
XX	XX	Example 13; SEQ ID NO 4; 60pp; English.		
PS	XX			
CC	XX	The invention relates to a novel compound comprising a sequence targeted to a nucleic acid encoding G protein-coupled receptor (GPCR) 49, that specifically hybridises with the nucleic acid encoding G protein-coupled receptor 49 and inhibits its expression. The compound of the invention demonstrates neuroprotective activity and may be useful for preparing a composition for treating neural or endocrine system disorders, as well as during gene and antisense therapy. The current sequence is that of the human G protein-coupled receptor (GPCR) 49 DNA of the invention.		
CC	XX			
QQ	SQ	Sequence 2724 BP; 665 A; 723 C; 554 G; 782 T; 0 U; 0 Other;		
 Query Match 100.0%; Score 2724; DB 12; Length 2724; Best Local Similarity 100.0%; Pred. No. 0; Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	1	ATGACACCTCCCGCTCGGTGCTCTCTCTGCTGCTGCTGCAGCTGGGAC	60	
Dd	1	ATGACACCTCCCGCTCGGTGCTCTCTCTGCTGCTGCTGCAGCTGGGAC	60	
Qy	61	GGGGGAGCTCTCCCAAGTCTGTGTGCTGAGGGGCTGCCCAACA	120	

Db 61 GGGGGCAGCTCTCCAGGCTCTGGTGTCTGAGGGGCTGCCCCACACACTGTCTATGCG 120
Qy 121 GAGCCCGACGCGAGGATGTCTCAGGGTGGAGTCTCTCCGACCTGGGGCTCTCGAGCTG 180
Db 121 GAGCCCGACGCGAGGATGTCTCAGGGTGGAGTCTCTCCGACCTGGGGCTCTCGAGCTG 180
Qy 181 CCTTCCAACTCAGGGTCTTCACTCTCCCTACCTAGACCTCAGTATGAAACAATCAGTCA 240
Db 181 CCTTCCAACTCAGGGTCTTCACTCTCCCTACCTAGACCTCAGTATGAAACAATCAGTCA 240
Qy 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTCTCCCTGAGGAGTACGTCTTGGGGAAC 300
Db 241 CTGCTCCCGAATCCCTCCAGTCTCCGCTCTCCCTGAGGAGTACGTCTTGGGGAAC 300
Qy 301 GCTCTGACATACATTTCCCAAGGAGCAATCTACCTGGCTTTTACAGTCTTTAAAGTCTTATG 360
Db 301 GCTCTGACATACATTTCCCAAGGAGCAATCTACCTGGCTTTTACAGTCTTTAAAGTCTTATG 360
Qy 361 CTGCAGATTAATCAGCTAAGCACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db 361 CTGCAGATTAATCAGCTAAGCACGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT 420
Qy 421 CAATCCCTGCTCTGGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAAGTGGC 480
Db 421 CAATCCCTGCTCTGGATGCTAAACACATCAGCTATGTCGCCCAAGCTGTTTCAAGTGGC 480
Qy 481 CTGCATTTCCCTGAGGACCTCTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGTCCAG 540
Db 481 CTGCATTTCCCTGAGGACCTCTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGTCCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCAATGCGCAATGCGCAATGCGCAATGCGCAATGCGCAATGCGCAAT 600
Db 541 GCTTTTGAAGTTTATCGGCAATGCGCAATGCGCAATGCGCAATGCGCAATGCGCAATGCGCAAT 600
Qy 601 ATACCAAGTATGCTTTGGAAACCTCTCCAGCTTGGTGTAGTCTTACATCTCCATAACAAT 660
Db 601 ATACCAAGTATGCTTTGGAAACCTCTCCAGCTTGGTGTAGTCTTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 721 TTAATTTACAATTAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTTAA 780
Db 721 TTAATTTACAATTAACCTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTTAA 780
Qy 781 GAACTAGATTTATAGCAACAATATCAGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGATTTATAGCAACAATATCAGTTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Qy 841 CCTTCTCTTATTAACAATATCTATGACAAATCCCAATCCCAATTTGTTGGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATATCTATGACAAATCCCAATCCCAATTTGTTGGGAGATCTGCT 900
Qy 901 TTTCAACAATTTACCTGAACCTAAGAACACTGACTCTGAAATGGTGGCTCAACAATAACTGAA 960
Db 901 TTTCAACAATTTACCTGAACCTAAGAACACTGACTCTGAAATGGTGGCTCAACAATAACTGAA 960
Qy 961 TTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAAGATC 1020
Db 961 TTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCAAGATC 1020
Qy 1021 TCATCTCTTCTCAAACTGCTGCAATCAGTTTACCTAATCTCCAAAGTCTAGATCTGTCT 1080
Db 1021 TCATCTCTTCTCAAACTGCTGCAATCAGTTTACCTAATCTCCAAAGTCTAGATCTGTCT 1080
Qy 1081 TACAACCTTATTAAGATTTTACCAAGTCTTTCAGTCTGCGCAAAAGCTTTCAGAAATTTGAC 1140
Db 1081 TACAACCTTATTAAGATTTTACCAAGTCTTTCAGTCTGCGCAAAAGCTTTCAGAAATTTGAC 1140
Qy 1141 CTAAGACATTAATGAATCTAGAAATTAAGTTGACATTTCCAGCAGTCTTCTAGCCTC 1200
Db 1141 CTAAGACATTAATGAATCTAGAAATTAAGTTGACATTTCCAGCAGTCTTCTAGCCTC 1200

Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTCTATTTATTTCAACCCCAATGCAATTTTCCA 1260
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTCTATTTATTTCAACCCCAATGCAATTTTCCA 1260
Qy 1261 TTGCGCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTTAACT 1320
Db 1261 TTGCGCATCCCTTAATAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTTAACT 1320
Qy 1321 GGGTTACATGTTTAACTCACTTAAATTTAAACAGAAATCAATGCTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGTTTAACTCACTTAAATTTAAACAGAAATCAATGCTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAAATTTTCCAGAACTCAAGTTTAAAGAAATGCTTATGCTTACAGTGTGT 1440
Db 1381 TCATCTGAAATTTTCCAGAACTCAAGTTTAAAGAAATGCTTATGCTTACAGTGTGT 1440
Qy 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGAAATTTCTAATCAATGGAATAAAGGTGCAAC 1500
Db 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGAAATTTCTAATCAATGGAATAAAGGTGCAAC 1500
Qy 1501 AGCAGTATGAGACCTTTCAAGAAAGATGCTCGAAATGTTTCAAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGAGACCTTTCAAGAAAGATGCTCGAAATGTTTCAAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAAGCCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAAGCCTGAAAGCCCTTCAATTCAGTGCAG 1620
Qy 1621 TGTTCACTTTCCCGAGGCCCCCTTCAAACTGCTGAAACCTGCTGAGTGGCTGGCTGATC 1680
Db 1621 TGTTCACTTTCCCGAGGCCCCCTTCAAACTGCTGAAACCTGCTGAGTGGCTGGCTGATC 1680
Qy 1681 AGAATGAGTGTGGAACCATAGCAGATTTCTGGCACTTACTTGAATGCTTTGGTGAATTTCA 1740
Db 1681 AGAATGAGTGTGGAACCATAGCAGATTTCTGGCACTTACTTGAATGCTTTGGTGAATTTCA 1740
Qy 1741 ACAGTTTTCAGATCCCCTCTGTATCAATTTCCCAATTTAACTGTTAAATTTGGGGTCAATGCA 1800
Db 1741 ACAGTTTTCAGATCCCCTCTGTATCAATTTCCCAATTTAACTGTTAAATTTGGGGTCAATGCA 1800
Qy 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCTGCTGGCTGGCTGGTGGATGCGTCTCACT 1860
Db 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCTGCTGGCTGGCTGGTGGATGCGTCTCACT 1860
Qy 1861 TTTGGCAGCTTTGCAACACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGTGG 1920
Db 1861 TTTGGCAGCTTTGCAACACATGCTGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGTGG 1920
Qy 1921 GGTGTTTGTGTCATTTTGTGCTTCAAGATTCATCTGTTTTTCTGCTTCTGCTGGCAGCCCTG 1980
Db 1921 GGTGTTTGTGTCATTTTGTGCTTCAAGATTCATCTGTTTTTCTGCTTCTGCTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTTGAAGAAAGCTTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTTGAAGAAAGCTTCCATTTTCTAGC 2040
Qy 2041 CTGAAAGTAACTATTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 CTGAAAGTAACTATTTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2101 CTGGGTGGCAGAAATGAGTATGGGCTCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGAGGCC 2160
Db 2101 CTGGGTGGCAGAAATGAGTATGGGCTCTCCCTCTCTGCTGCTGCTTTCCTTTTGGGAGGCC 2160
Qy 2161 AGCAACATGGGCTACATGCTGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 2161 AGCAACATGGGCTACATGCTGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Qy 2221 ACCATTGCTACCAAGCTCTACTGCAATTTTGGCAAGGAGACCTGGAGAAATTTTGG 2280
Db 2221 ACCATTGCTACCAAGCTCTACTGCAATTTTGGCAAGGAGACCTGGAGAAATTTTGG 2280

Qy	721	TTAAATTACAATAACCTTGTGATGAATTTCCCACTGCAATTAGGACACTCTCTCAACCTTTAAA	780
Db	769	TTAAATTACAATAACCTTGTGATGAATTTCCCACTGCAATTAGGACACTCTCTCAACCTTTAAA	828
Qy	781	GAACCTAGGATTTTCATAGCAACAATATCAGCTGCAGTACCTCAGAAAGACATTTGTAGGCAAC	840
Db	829	GAACCTAGGATTTTCATAGCAACAATATCAGCTGCAGTACCTCAGAAAGACATTTGTAGGCAAC	888
Qy	841	CTTCTCTTATTACAATACATTTTCTATGACAATCCCATCCAAATTTGTGGGAGATCTGCT	900
Db	889	CTTCTCTTATTACAATACATTTTCTATGACAATCCCATCCAAATTTGTGGGAGATCTGCT	948
Qy	901	TTTCAACATTTTACCTGAACTTAAGAACACTGACTCTGAAATGGTGGCTCACAATAACTGAA	960
Db	949	TTTCAACATTTTACCTGAACTTAAGAACACTGACTCTGAAATGGTGGCTCACAATAACTGAA	1008
Qy	961	TTTCTCTGATTTAACTCTGAACTTGCAAACTCGAGAGTCTGACTTTTAACTGAGACACAGATC	1020
Db	1009	TTTCTCTGATTTAACTCTGAACTTGCAAACTCGAGAGTCTGACTTTTAACTGAGACACAGATC	1068
Qy	1021	TCATCTCTTCTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGCTAGATCTGTCT	1080
Db	1069	TCATCTCTTCTCTCAAAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTGCTAGATCTGTCT	1128
Qy	1081	TACAACTTATAGNAGATTTTACCAGTHTTTTTCAGTCTGCCAAAAGCTTTCAGAAAATTGAC	1140
Db	1129	TACAACTTATAGNAGATTTTACCAGTHTTTTTCAGTCTGCCAAAAGCTTTCAGAAAATTGAC	1188
Qy	1141	CTAAGACATAAATGAAATCTACGAAATTAAGTTTGACACTTTTCCAGCAGTTGCTTAGGCCTC	1200
Db	1189	CTAAGACATAAATGAAATCTACGAAATTAAGTTTGACACTTTTCCAGCAGTTGCTTAGGCCTC	1248
Qy	1201	CGATCGCTGAAATTTGGCTTGGAACAAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT	1260
Db	1249	CGATCGCTGAAATTTGGCTTGGAACAAAATTTGCTATTATTCACCCCAATGCAATTTTCCACT	1308
Qy	1261	TTGGCATCCCTAATAAGCTGGAGCTATCGTCCAACTCTGTCGTGCTTTTCCCTATAACT	1320
Db	1309	TTGGCATCCCTAATAAGCTGGAGCTATCGTCCAACTCTGTCGTGCTTTTCCCTATAACT	1368
Qy	1321	GGGTACATCGTTTAACTCACTTAAATAATAACAGGAAATCATGCTTTACAGAGCTTGATA	1380
Db	1369	GGGTACATCGTTTAACTCACTTAAATAATAACAGGAAATCATGCTTTACAGAGCTTGATA	1428
Qy	1381	TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGACCTTACCAAGTGTCTGT	1440
Db	1429	TCATCTGAAAACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGACCTTACCAAGTGTCTGT	1488
Qy	1441	GCATTTGGAGTGTGAGAGTGCCTATAAGATTTTCTAATCAATGGAATAAGGTGACAC	1500
Db	1489	GCATTTGGAGTGTGAGAGTGCCTATAAGATTTTCTAATCAATGGAATAAGGTGACAC	1548
Qy	1501	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1560
Db	1549	AGCAGTATGGACGACCTTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT	1608
Qy	1561	GACCTTGAAGATTTCTGCTTTGACTTTTGAAGAGAGACCTGAAAGCCCTTCATTCAAGTGCAG	1620
Db	1609	GACCTTGAAGATTTCTGCTTTGACTTTTGAAGAGAGACCTGAAAGCCCTTCATTCAAGTGCAG	1668
Qy	1621	TGTTTCACTTTCCCGAGGCCCTTTCAAACCTGTGTGAACACCTGCTTGATGCTCGCTGATC	1680
Db	1669	TGTTTCACTTTCCCGAGGCCCTTTCAAACCTGTGTGAACACCTGCTTGATGCTCGCTGATC	1728
Qy	1681	AGAATTGGAGTGTGGACCATAGCAGTTTCTGGCACTTACTTGTAAATGCTTTTGGTGCATTTCA	1740
Db	1729	AGAATTGGAGTGTGGACCATAGCAGTTTCTGGCACTTACTTGTAAATGCTTTTGGTGCATTTCA	1788
Qy	1741	ACAGTTTTTCAGATCCCTCTGTGACATTTTCCCGCAATTAACATGTTTAATTTGGGGTCATCGCA	1800
Db	1789	ACAGTTTTTCAGATCCCTCTGTGACATTTTCCCGCAATTAACATGTTTAATTTGGGGTCATCGCA	1848

RESULT 8
ADQ80249
ID ADQ
XX
AC ADQ
XX
DT 21-6
XX

RESULT 12
AAX23980
ID AAX23980 standard; DNA; 4558 BP.
XX AC
XX AC AAX23980;
XX DT 25-JUN-1999 (first entry)
XX DE Human HG38 DNA.
XX KW HG38; human; G-protein coupled glycoprotein hormone receptor; brain;
XX KW endocrine system; skeletal muscle; spinal cord; placenta; development;
XX KW receptor activity modulator; ss.
XX OS Homo sapiens.
XX FN WO9915660-A1.
XX PD 01-APR-1999.
XX PF 24-SEP-1998; 98WO-US019979.
XX PR 24-SEP-1997; 97US-0059863P.
XX PA (MERI) MERCK & CO INC.
XX PI Liu Q, Bailey WJ, McDonald TP;
XX DR WPI; 1999-254711/21.
XX DR P-PSDB; AA093889.
XX PT Human G-protein coupled glycoprotein hormone receptor HG38.
XX PS Claim 4; Fig 1A-C; 74pp; English.
CC This invention describes a novel human G-protein coupled glycoprotein
CC hormone receptor, HG38. Glycoprotein hormone receptors are important in
CC the endocrine system and HG38 may be involved in development and function
CC of the skeletal muscle, spinal cord, placenta and to a lesser extent, the
CC brain. The transgenic animal may be useful for studying tissue and
CC temporal specific expression or activity of the HG38 receptor, as well as
CC for studying the ability of a variety of compounds to act as modulators
CC of HG38 receptor activity
SQ Sequence 4558 BP; 1225 A; 1063 C; 914 G; 1355 T; 0 U; 1 Other;
Query Match 100.0%; Score 2724; DB 2; Length 4558;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGACACACTCCCGCTCGGTGTGCTCCTGTCTTCCTGGCTGTGCTGCAGTGGCGACC 60
Dd 249 ATGACACACTCCCGCTCGGTGTGCTCCTGTCTTCCTGGCTGTGCTGCAGTGGCGACC 308
Qy 61 GGGGGCAGCTCTCCAGGTCTGGTGTTGTGTGAGGGGCTGCCCACACATGTCAATTGC 120
Dd 309 GGGGGCAGCTCTCCAGGTCTGGTGTTGTGTGAGGGGCTGCCCACACATGTCAATTGC 368
Qy 121 GAGCCCGACGCGAGGATTTCTCAGGGTGAAGTCTCCGACCTGGGGCTCTCGAGCTG 180
Dd 369 GAGCCCGACGCGAGGATTTGTTCTCAGGGTGAAGTCTCCGACCTGGGGCTCTCGAGCTG 428
Qy 181 CCTTCCAACCTCAGCGTCTTCACCTCTCCCTACCTCAGTATGAAACAACATCATGTCAG 240
Dd 429 CCTTCCAACCTCAGCGTCTTCACCTCTCCCTACCTCAGTATGAAACAACATCATGTCAG 488
Qy 241 CTGCTCCCGAATCCCCTGCCAGTCTCCGCTTCCTGGAGGAGTTACGTCCTTCGCGGAAC 300
Dd 489 CTGCTCCCGAATCCCCTGCCAGTCTCCGCTTCCTGGAGGAGTTACGTCCTTCGCGGAAC 548
Qy 301 GCTCTGACATACATTCACAGGAGCATCTACCTGGCTTTTACAGTCTTAAGTTCCTTAG 360
Dd 549 GCTCTGACATACATTCACAGGAGCATCTACCTGGCTTTTACAGTCTTAAGTTCCTTAG 608

1441 GCATTTGGAGTGTGAGAAAGCCCTATAAGATTTCTTAATCAATGAATATAAGGTGACAAC 1500
Db |||||
1489 GCATTTGGAGTGTGAGAAAGCCCTATAAGATTTCTTAATCAATGAATATAAGGTGACAAC 1748
Qy |||||
1501 AGCAGTATGAGACGACCTTTATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1560
Db |||||
1749 AGCAGTATGAGACGACCTTTATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGT 1808
Qy |||||
1561 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db |||||
1809 GACCTTGAAGATTTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1868
Qy |||||
1621 TGTTCACCTTCCACAGGCCCTTCAAAACCCCTGTGAACACCTGCTTGTAGTGGCTGGCTGATC 1680
Db |||||
1869 TGTTCACCTTCCACAGGCCCTTCAAAACCCCTGTGAACACCTGCTTGTAGTGGCTGGCTGATC 1928
Qy |||||
1681 AGAATTGAGTGTGAGACCATAGCAGTCTCTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCA 1740
Db |||||
1929 AGAATTGAGTGTGAGACCATAGCAGTCTCTGGCAGCTTACTTGTAAATGCTTTGGTGACTTCA 1988
Qy |||||
1741 ACAGTTTTCAGATCCCTCTGTGATCAATTTCCCCCAATTAACCTGTTAAATTTGGGGTCAATGCA 1800
Db |||||
1989 ACAGTTTTCAGATCCCTCTGTGATCAATTTCCCCCAATTAACCTGTTAAATTTGGGGTCAATGCA 2048
Qy |||||
1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCGCTGCTGGTGTGGTGTGGATGCGTTCACT 1860
Db |||||
2049 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCGCTGCTGGTGTGGTGTGGATGCGTTCACT 2108
Qy |||||
1861 TTTGGCAGCTTTGACACGACATGGTGGTGGTGGGAGAAATGGGGTGTGGTGCATGTCATT 1920
Db |||||
2109 TTTGGCAGCTTTGACACGACATGGTGGTGGTGGGAGAAATGGGGTGTGGTGCATGTCATT 2168
Qy |||||
1921 GGTTTTGTGTCATTTTGTCTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTG 1980
Db |||||
2169 GGTTTTGTGTCATTTTGTCTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTG 2228
Qy |||||
1981 GAGCGTGGTCTCTGCAAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db |||||
2229 GAGCGTGGTCTCTGCAAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2288
Qy |||||
2041 CTGAAGTAATCATTTTGTCTCTGTGCCCTGTGCGCTTGACCATGGCGCAGTTTCCCTG 2100
Db |||||
2289 CTGAAGTAATCATTTTGTCTCTGTGCCCTGTGCGCTTGACCATGGCGCAGTTTCCCTG 2348
Qy |||||
2101 CTGGTGGCAGGAATGAGGCGCTCCCTCTCTGCGCTGCTTGGCCCTTTTGGGGAGCCC 2160
Db |||||
2349 CTGGTGGCAGGAATGAGGCGCTCCCTCTCTGCGCTGCTTGGCCCTTTTGGGGAGCCC 2408
Qy |||||
2161 AGCACCATGGGCTACATGCTCGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2220
Db |||||
2409 AGCACCATGGGCTACATGCTCGCTCTCATCTTGTCTCAATTCCTTTGCTTCTCTCATGATG 2468
Qy |||||
2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGAACAAGGAGAGACCTGGAGAATATTTGG 2280
Db |||||
2469 ACCATTGCTACACCAAGCTCTACTGCAATTTGGAACAAGGAGAGACCTGGAGAATATTTGG 2528
Qy |||||
2281 GACTGCTCTATGTTAAACACATTTGCCCTGTGCTCTTCAACCACTGATCCTTAACTGC 2340
Db |||||
2529 GACTGCTCTATGTTAAACACATTTGCCCTGTGCTCTTCAACCACTGATCCTTAACTGC 2588
Qy |||||
2341 CCTGGGCTTCTTGTGCTCTCTCTCTTTAATAAACCTTACATTTATCAGTCCCTGAAGTA 2400
Db |||||
2589 CCTGGGCTTCTTGTGCTCTCTCTCTTTAATAAACCTTACATTTATCAGTCCCTGAAGTA 2648
Qy |||||
2401 ATTAAGTTTATCTCTCTGGTGGTAGTCCCACTTCTCGATGCTCTCAATCCCTCTCTTAC 2460
Db |||||
2649 ATTAAGTTTATCTCTCTGGTGGTAGTCCCACTTCTCGATGCTCTCAATCCCTCTCTTAC 2708
Qy |||||
2461 ATCTTGTTCATCTCTCACTTTAAGAGGATCTGTGAGCCTGAGAAACCAACCTACGTC 2520
Db |||||
2709 ATCTTGTTCATCTCTCACTTTAAGAGGATCTGTGAGCCTGAGAAACCAACCTACGTC 2768
Qy |||||
2521 TGGACAAGATCAAAACACCAAGCTTGTATGTCATTTAACTCTGTATGATGTGCAAAAACAG 2580

Db |||||
2769 TGGACAAGATCAAAACACCAAGCTTGTATGTCATTTAACTCTGTATGATGTGCAAAAACAG 2828
Qy |||||
2581 TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTTACAGCTCCAGCATCACATTATGACCTG 2640
Db |||||
2829 TCCTGTGACTCAACTCAAGCCTTGGTAAACCTTTTACAGCTCCAGCATCACATTATGACCTG 2888
Qy |||||
2641 CCTCCAGTTCCGTCGTCATCACAGCTTATCCAGTGAATCCAGTGAATCCATCTTCTCTCT 2700
Db |||||
2889 CCTCCAGTTCCGTCGTCATCACAGCTTATCCAGTGAATCCAGTGAATCCATCTTCTCTCT 2948
Qy |||||
2701 GTGGCATTTGTGCCATGTCCTCTAA 2724
Db |||||
2949 GTGGCATTTGTGCCATGTCCTCTAA 2972
RESULT 13
ADL12472
ID ADL12472 standard; cDNA; 4570 BP.
XX
AC ADL12472;
XX
DT 06-MAY-2004 (first entry)
XX
DB Human steroid-induced C3A liver cell cDNA #201.
XX
KW ss; gene; Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX
OS Homo sapiens.
XX
PN US6673549-B1.
XX
PD 06-JAN-2004.
XX
PF 12-OCT-2001; 2001US-00976594.
XX
PR 12-OCT-2000; 2000US-0240409P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Furness LM, Buchbinder JL;
XX
DR WPI; 2004-068610/07.
XX
CC Combination useful for preparing a composition for treating liver
CC disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
CC comprises cDNAs that are differentially expressed in response to steroid
CC treatment.
XX
PS Claim 1; SEQ ID NO 201; 141pp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human cDNA which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 4570 BP; 1228 A; 1069 C; 918 G; 1352 T; 0 U; 3 Other;

[illegible]

Db 481 CTGCATTCCTGAGGCACCTGTGGCTGGATGACAAATGCGTTAAACGAATATCCCGTCCAG 540
Qy 541 GCTTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
Db 541 GCTTTTAGAAGTTTATCGGCATTGCAAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
Qy 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTGGTAGTTCTACATCTCCATAACAAT 660
Db 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTGGTAGTTCTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 721 TTAATAACAATAACCTTGATGAATTCGCCACTGCAATTAGGACACTCTCAACCTTAAA 780
Db 721 TTAATAACAATAACCTTGATGAATTCGCCACTGCAATTAGGACACTCTCAACCTTAAA 780
Qy 781 GAACTAGGATTCATAGCAACAAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAC 840
Db 781 GAACTAGGATTCATAGCAACAAATATCAGGTGCGATACCTGAGAAAGCAATTTGTAGGCAC 840
Qy 841 CCTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 841 CCTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Qy 901 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTGAAATGGTGCCCTCAAAATACGTAA 960
Db 901 TTTCAACATTTACCTGAACCTAAGAACACTGACTCTGAAATGGTGCCCTCAAAATACGTAA 960
Qy 961 TTTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGGACAGATC 1020
Db 961 TTTCTGATTTAACTGGAACTGCAACCTGGAGAGTCTGACTTTAACTGGAGGACAGATC 1020
Qy 1021 TCATCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
Db 1021 TCATCTCTCTCAAAACCGTCTGCAATCAGTTTACCTAATCTCCAAGTCTAGATCTGTCT 1080
Qy 1081 TACAACCTATTAGAGATTTACCCAGTTTTTTCAGTCTGCAAAAGCTTTAGAAAATTTGAC 1140
Db 1081 TACAACCTATTAGAGATTTACCCAGTTTTTTCAGTCTGCAAAAGCTTTAGAAAATTTGAC 1140
Qy 1141 CTAAGACATAATGAAATCTAGAAATTTAAAGTTGACACTTTCCAGCAGTTCCTTAGCCTC 1200
Db 1141 CTAAGACATAATGAAATCTAGAAATTTAAAGTTGACACTTTCCAGCAGTTCCTTAGCCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATGCTATTATTCACCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATGCTATTATTCACCCCAATGCAATTTTCCACT 1260
Qy 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTGTGCTCTTTTCTCTATAACT 1320
Db 1261 TTGCCATCCCTAATAAAGCTGGACCTATCGTCCAACTCTGTGCTCTTTTCTCTATAACT 1320
Qy 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTCTGT 1440
Db 1381 TCATCTGAATTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTCTGT 1440
Qy 1441 GCAATTTGAGTGTGTGAGAAATGCTTATAGAAATTTCTAATCAATGGAATAAAGGTGCAAC 1500
Db 1441 GCAATTTGAGTGTGTGAGAAATGCTTATAGAAATTTCTAATCAATGGAATAAAGGTGCAAC 1500
Qy 1501 AGCAGTATGAGCGACTTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGAGCGACTTTCATAAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTGAGATTTTCTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1561 GACCTTGAGATTTTCTCTGCTTGAATTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620

Qy 1621 TGTTCACTTCCCAGAGGCCCTTCAAACCTGTGTGAAACACCTGTGTGATGGCTGTGCTGATC 1680
Db 1621 TGTTCACTTCCCAGAGGCCCTTCAAACCTGTGTGAAACACCTGTGTGATGGCTGTGCTGATC 1680
Qy 1681 AGAATTTGGAGTGTGGACCATAGCAGTTCTGSCACTTACTTGTAAATGCTTTTGGTGAATTTCA 1740
Db 1681 AGAATTTGGAGTGTGGACCATAGCAGTTCTGSCACTTACTTGTAAATGCTTTTGGTGAATTTCA 1740
Qy 1741 ACAGTTTTTCAGATCCCCTCTGTACATTTCCCCCAATTAACCTGTAAATTTGGGGTCAATGCA 1800
Db 1741 ACAGTTTTTCAGATCCCCTCTGTACATTTCCCCCAATTAACCTGTAAATTTGGGGTCAATGCA 1800
Qy 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGTGCTGGTGTGGATGCGTTCACT 1860
Db 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTGCCTGTGCTGGTGTGGATGCGTTCACT 1860
Qy 1861 TTTGGCAGCTTTGACGACATGCTGTGGTGGGAGATGGGGTTGGTGGCCATGCTCAAT 1920
Db 1861 TTTGGCAGCTTTGACGACATGCTGTGGTGGGAGATGGGGTTGGTGGCCATGCTCAAT 1920
Qy 1921 GGTTTTTTGTCCATTTTTTGTCTTCAAGATCATCTGTTTTTCTGTCTTACTCTGGCAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTTTGTCTTCAAGATCATCTGTTTTTCTGTCTTACTCTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTTCCATTTTCTAGC 2040
Qy 2041 CTGAAATGATATTTTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2100
Db 2041 CTGAAATGATATTTTTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2100
Qy 2101 CTGGGTGGCAGAGTATGGCGCTCCCTCTCTGCTGCTGCTTTGGGAGAGCC 2160
Db 2101 CTGGGTGGCAGAGTATGGCGCTCCCTCTCTGCTGCTGCTTTGGGAGAGCC 2160
Qy 2161 AGCAACATGGGCTACATGGTGTCTCATCTGTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Db 2161 AGCAACATGGGCTACATGGTGTCTCATCTGTGCTCAATTTCCCTTTGCTTCTCATGATG 2220
Qy 2221 ACCATGCTTACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATTTTGG 2280
Db 2221 ACCATGCTTACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATTTTGG 2280
Qy 2281 GACTGCTTATGTTAAACACATTTGCCCTGTGCTTCTTACCAACTGCATCTTAAACTGC 2340
Db 2281 GACTGCTTATGTTAAACACATTTGCCCTGTGCTTCTTACCAACTGCATCTTAAACTGC 2340
Qy 2341 CCTGTGCTTTTGTCTTCTCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Db 2341 CCTGTGCTTTTGTCTTCTCTCTTTAATAAACCTTACATTTATCAGTCTCTGAAGTA 2400
Qy 2401 ATTAAGTTTATCTTCTGTGGTGTAGTCCACTTCTGCTGATGCTCAATCCCTTCTCTAC 2460
Db 2401 ATTAAGTTTATCTTCTGTGGTGTAGTCCACTTCTGCTGATGCTCAATCCCTTCTCTAC 2460
Qy 2461 ATCTTTTCAATCTCTCACTTTAAGAGGATCTGGTGTAGGAGCAAACTTACGTC 2520
Db 2461 ATCTTTTCAATCTCTCACTTTAAGAGGATCTGGTGTAGGAGCAAACTTACGTC 2520
Qy 2521 TGGACAAGATCAAAACCCCAAGCTTGTATGTCAATTAACCTCTGTATGTGTCGAAAAACAG 2580
Db 2521 TGGACAAGATCAAAACCCCAAGCTTGTATGTCAATTAACCTCTGTATGTGTCGAAAAACAG 2580
Qy 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTACAGCTCCAGCATCTTATGACCTG 2640
Db 2581 TCCTGTGACTCAACTCAAGCCTTGGTAACTTTACAGCTCCAGCATCTTATGACCTG 2640
Qy 2641 CCTCCAGTTCGCGCCATCAACAGCTTATCCAGTGTAGAGCTGCGATCTTTTCTCT 2700
Db 2641 CCTCCAGTTCGCGCCATCAACAGCTTATCCAGTGTAGAGCTGCGATCTTTTCTCT 2700

QY 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACCTTCCAGCAGTTGCTTAGCCCTC 1200
DB 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACCTTCCAGCAGTTGCTTAGCCCTC 1200
QY 1201 CGATCGCTGAATTTGGCTTGGACAAATTTGCTATTATTTACCCCAATGCAATTTCCCACT 1260
DB 1201 CGATCGCTGAATTTGGCTTGGACAAATTTGCTATTATTTACCCCAATGCAATTTCCCACT 1260
QY 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCGTCTGCTTTTCCCTATAACT 1320
DB 1261 TTGCCATCCCTAATAAGCTGGACCTATCGTCAACCTCGTCTGCTTTTCCCTATAACT 1320
QY 1321 GGGTTACATGGTTAACTCACTTAAAAATTAAACAGGAATCATGCTTACAGAGCTTGATA 1380
DB 1321 GGGTTACATGGTTAACTCACTTAAAAATTAAACAGGAATCATGCTTACAGAGCTTGATA 1380
QY 1381 TCATCTGAAAACTTTCCGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACCAAGTGTGT 1440
DB 1381 TCATCTGAAAACTTTCCGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACCAAGTGTGT 1440
QY 1441 GCATTTGGAGTGTGAGAAATCCCTATAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
DB 1441 GCATTTGGAGTGTGAGAAATCCCTATAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
QY 1501 AGCAGTATGGACGACCTTCATAGAAAGATGCTGGAATTTTCAAGCTCAAGATGAACGT 1560
DB 1501 AGCAGTATGGACGACCTTCATAGAAAGATGCTGGAATTTTCAAGCTCAAGATGAACGT 1560
QY 1561 GACCTTGAAGATTTCTGCTTACCTTGAAGGAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
DB 1561 GACCTTGAAGATTTCTGCTTACCTTGAAGGAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
QY 1621 TGTTCACCTTCCCGAGGCCCTTCAAACTGCTGAAACACCTGCTTGAAGGCTGCTGATC 1680
DB 1621 TGTTCACCTTCCCGAGGCCCTTCAAACTGCTGAAACACCTGCTTGAAGGCTGCTGATC 1680
QY 1681 AGAATGGAGTGTGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTGGTGACTTCA 1740
DB 1681 AGAATGGAGTGTGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTGGTGACTTCA 1740
QY 1741 ACAGTTTTTCCAGATCCCTCTGACATTTCCCCCATTAACCTTAATTTGGGGTCAACGCA 1800
DB 1741 ACAGTTTTTCCAGATCCCTCTGACATTTCCCCCATTAACCTTAATTTGGGGTCAACGCA 1800
QY 1801 GCAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGCTGGCTGCTGCTGCTGCTGCTGCT 1860
DB 1801 GCAGTGAACATGCTCAACGGAGTCTCCAGTGCCTGCTGGCTGCTGCTGCTGCTGCTGCT 1860
QY 1861 TTTGGCAGCTTTGCAACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCATT 1920
DB 1861 TTTGGCAGCTTTGCAACATGGTGCCTGGTGGGAGAAATGGGGTTGGTTGCCATGTCATT 1920
QY 1921 GGTTTTTTGTCCATTTTGTCTCAGAAATCATGTTTTTCCCTGCTTACTCTGGCAGCCCTG 1980
DB 1921 GGTTTTTTGTCCATTTTGTCTCAGAAATCATGTTTTTCCCTGCTTACTCTGGCAGCCCTG 1980
QY 1981 GAGCGTGGGTCTCTGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB 1981 GAGCGTGGGTCTCTGTGAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
QY 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
DB 2041 CTGAAAGTAATCATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
QY 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
DB 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
QY 2161 AGCACCATGGGCTACATGCTGCTCTCATCTTGTGCTCAATTCCTTTGCTTCTCATGATG 2220
DB 2161 AGCACCATGGGCTACATGCTGCTCTCATCTTGTGCTCAATTCCTTTGCTTCTCATGATG 2220

QY 2221 ACCATTGCTTACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTTGG 2280
DB 2221 ACCATTGCTTACCAAGCTCTACTGCAATTTTGGACAAGGGAGACCTGGAGAAATATTTGG 2280
QY 2281 GACTGCTCTATGGTAAAAACACATTGGCCCTGTTGCTCTTCCACCACTGCAATCTAAACTGC 2340
DB 2281 GACTGCTCTATGGTAAAAACACATTGGCCCTGTTGCTCTTCCACCACTGCAATCTAAACTGC 2340
QY 2341 CCTGGGCTTTCTTGTCT 2400
DB 2341 CCTGGGCTTTCTTGTCT 2400
QY 2401 ATTAAGTTTATCT 2460
DB 2401 ATTAAGTTTATCT 2460
QY 2461 ATCTTGTTCATCT 2520
DB 2461 ATCTTGTTCATCT 2520
QY 2521 TGGACAAGATCAAAACACCCCAAGCTTGTGATGTCATTTAACTCTGATGATGTCGAAAAACAG 2580
DB 2521 TGGACAAGATCAAAACACCCCAAGCTTGTGATGTCATTTAACTCTGATGATGTCGAAAAACAG 2580
QY 2581 TCCTGTGACTCAACTCAAGCCCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
DB 2581 TCCTGTGACTCAACTCAAGCCCTTGGTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
QY 2641 CCTCCAGTTCCTGCTCCATCAGCTTTATCCAGTACTGAGAGCTGCGCATCTTTCTCTCT 2700
DB 2641 CCTCCAGTTCCTGCTCCATCAGCTTTATCCAGTACTGAGAGCTGCGCATCTTTCTCTCT 2700
QY 2701 GTGGCAATTTGTCCCATGTCCTCTAA 2724
DB 2701 GTGGCAATTTGTCCCATGTCCTCTAA 2724

Search completed: July 12, 2005, 14:17:31
Job time : 1585.87 secs

This page Blank (uspto)

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2724	100.0	2724	4	US-09-170-496D-263	Sequence 263, Appl
2	2724	100.0	4570	4	US-09-976-594-201	Sequence 201, Appl
3	2720.8	99.9	2724	4	US-09-170-496D-277	Sequence 277, Appl
4	710.4	26.1	723	4	US-09-495-050A-220	Sequence 220, Appl
5	556.4	20.4	2988	4	US-09-799-451-723	Sequence 723, Appl
6	540	19.8	2612	4	US-09-495-050A-214	Sequence 214, Appl
7	511.2	18.8	4203	2	US-08-866-757-1	Sequence 1, Appl
8	511.2	18.8	4203	3	US-09-153-593-1	Sequence 1, Appl
9	125.8	4.6	2703	4	US-09-482-273-75	Sequence 75, Appl
10	123.8	4.5	473	2	US-08-866-757-3	Sequence 3, Appl
11	123.8	4.5	473	3	US-09-153-593-3	Sequence 3, Appl
12	115.4	4.2	2709	4	US-09-482-273-101	Sequence 3, Appl
13	106.8	3.9	2413	4	US-09-016-434-1459	Sequence 1459, Appl
14	100.2	3.7	3710	3	US-07-741-453A-67	Sequence 67, Appl
15	100	3.7	4417	3	US-07-741-453A-52	Sequence 52, Appl
16	98.6	3.6	2292	4	US-09-826-509-402	Sequence 57, Appl
17	98.6	3.6	2469	4	US-09-186-350A-54	Sequence 402, Appl
18	97	3.6	2292	4	US-09-826-509-394	Sequence 54, Appl
19	97	3.6	2292	4	US-09-826-509-398	Sequence 394, Appl
20	97	3.6	2292	4	US-09-826-509-406	Sequence 398, Appl
21	97	3.6	2292	4	US-09-826-509-414	Sequence 406, Appl
22	95.4	3.5	2292	4	US-09-826-509-410	Sequence 414, Appl
23	93.8	3.4	2292	4	US-09-826-509-418	Sequence 410, Appl
24	93.8	3.4	3549	4	US-09-826-509-418	Sequence 418, Appl
25	92.2	3.4	2292	4	US-09-826-509-586	Sequence 586, Appl
26	91	3.3	2179	1	US-08-487-886-1	Sequence 422, Appl
27	91	3.3	2179	2	US-08-531-070A-1	Sequence 1, Appl

|||||
Db 361 CTGCAGAAATACAGCTAAGACACAGTACCCACAGAGCTCTGCGAATTTTCGGAAGCCTT 420
Qy 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Qy 481 CTGCATTCCTTGAGGACCTGTGGCTGGATGACAAATGGCTTTAAACAGAAATCCCGTCCAG 540
Db 481 CTGCATTCCTTGAGGACCTGTGGCTGGATGACAAATGGCTTTAAACAGAAATCCCGTCCAG 540
Qy 541 GCTTTTAGAAGTTTATCGGCATTCGACGCTGATGCTGGCTGGCTGGAACAAATAACACCAC 600
Db 541 GCTTTTAGAAGTTTATCGGCATTCGACGCTGATGCTGGCTGGCTGGAACAAATAACACCAC 600
Qy 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Db 601 ATACCAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Qy 721 TTAAATTTACAATTAACCTTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
Db 721 TTAAATTTACAATTAACCTTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
Qy 781 GAACTAGGATTTCTATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTCTATAGCAACAATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Qy 841 CCTCTCTTTATACAAATCTATGACAAATCCCATCCAAATTTGTGGAGATCTGCT 900
Db 841 CCTCTCTTTATACAAATCTATGACAAATCCCATCCAAATTTGTGGAGATCTGCT 900
Qy 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAATGTGCTCACAATAAATCTGAA 960
Db 901 TTTCAACATTTACCTGAACTAAGAACACTGACTCTGAATGTGCTCACAATAAATCTGAA 960
Qy 961 TTTCTGATTTAACTGGAACCTGCAACCTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
Db 961 TTTCTGATTTAACTGGAACCTGCAACCTGGAGAGTCTGACTTTAACTGGAGCAGATC 1020
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCATCAGTACCTAATCTCCAAAGTCTAGATCTGTCT 1080
Db 1021 TCATCTCTTCTCAAAACCGTCTGCATCAGTACCTAATCTCCAAAGTCTAGATCTGTCT 1080
Qy 1081 TACAACCTATTAGAGATTTTACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAAATTCAC 1140
Db 1081 TACAACCTATTAGAGATTTTACCAAGTTTTCAGTCTGCCAAAGCTTCAGAAAATTCAC 1140
Qy 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTGTCTTAGCCTC 1200
Db 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTGTCTTAGCCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
Qy 1261 TTGCCATCCCTTAAATAAGCTGACCTATGCTCCAACTCTGCTGCTCTTTTCTTAACT 1320
Db 1261 TTGCCATCCCTTAAATAAGCTGACCTATGCTCCAACTCTGCTGCTCTTTTCTTAACT 1320
Qy 1321 GGGTTTACATGGTTTAACTCACTTAAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Db 1321 GGGTTTACATGGTTTAACTCACTTAAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAAACTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGTGTGT 1440
Db 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGTGTGT 1440
Qy 1441 GCATTTGGAGTGTGTGAGAAATGCTCTATAGAAATTTCTAATCAATGGAATTAAGAGTGACAA 1500
|||||

Db 1441 GCATTTGGAGTGTGTGAGAAATGCCCTATAGAAATTTCTAATCAATGGAATTAAGAGTGACAA 1500
Qy 1501 AGCAGTATGACGACCTTTCATAGAAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGACGACCTTTCATAGAAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTCAAGATTTCCCTGCTTGTGCTTGTGAGGAAGACCTGAAAAGCCCTTCATTCAAGTGAC 1620
Db 1561 GACCTTCAAGATTTCCCTGCTTGTGCTTGTGAGGAAGACCTGAAAAGCCCTTCATTCAAGTGAC 1620
Qy 1621 TGTTCACCTTCCCAGAGCCCTTCAAAACCTCTGTGAAACACCTGTCTGTATGGCTGGCTGATC 1680
Db 1621 TGTTCACCTTCCCAGAGCCCTTCAAAACCTCTGTGAAACACCTGTCTGTATGGCTGGCTGATC 1680
Qy 1681 AGAATTCGAGTGTGGACATAGCAGATTTCTGGCACTTACTTGTGAATGCTTTGGTGACTTCA 1740
Db 1681 AGAATTCGAGTGTGGACATAGCAGATTTCTGGCACTTACTTGTGAATGCTTTGGTGACTTCA 1740
Qy 1741 ACAGTTCAGATCCCTCTGTACATTTCCCACTTAAACCTGTTAAATGGGGTTCATCGCA 1800
Db 1741 ACAGTTCAGATCCCTCTGTACATTTCCCACTTAAACCTGTTAAATGGGGTTCATCGCA 1800
Qy 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGCTGTGGTGTGGATGCTTCACT 1860
Db 1801 GCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGCTGGCTGTGGTGTGGATGCTTCACT 1860
Qy 1861 TTTGGCAGCTTTGCAACATGCTGCTGGTGGGAGAAATGGGGTGGTGGCATGTCAAT 1920
Db 1861 TTTGGCAGCTTTGCAACATGCTGCTGGTGGGAGAAATGGGGTGGTGGCATGTCAAT 1920
Qy 1921 GGTTCCTTTGTCATTTTGTTCAGAAATCATCTGCTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 1921 GGTTCCTTTGTCATTTTGTTCAGAAATCATCTGCTTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGTCTCTCTGTGAAATATTTCTGCAAAATTTGAAAACGAAAGCTCCATTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTCTGTGAAATATTTCTGCAAAATTTGAAAACGAAAGCTCCATTTCTAGC 2040
Qy 2041 CTGAAAGTAAATCAATTTTGTCTCTGTGCCCTGTGCTGCTTACCATGGCGGAGTTCCCTG 2100
Db 2041 CTGAAAGTAAATCAATTTTGTCTCTGTGCCCTGTGCTGCTTACCATGGCGGAGTTCCCTG 2100
Qy 2101 CTGGGTGGCAGAGTATGGCGCTCTCTGCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTG 2160
Db 2101 CTGGGTGGCAGAGTATGGCGCTCTCTGCTGCTGCTTCTGCTTCTGCTTCTGCTTCTGCTG 2160
Qy 2161 AGCACCATGGGCTACATGGTCTCATCTTGTCTCAATTTCCCTTCTGCTTCTCATGATG 2220
Db 2161 AGCACCATGGGCTACATGGTCTCATCTTGTCTCAATTTCCCTTCTGCTTCTCATGATG 2220
Qy 2221 ACATTTGCTTACCAAGCTCTACTGCAATTTTGGCAAGGGAGACCTGGAGAAATTTTGG 2280
Db 2221 ACATTTGCTTACCAAGCTCTACTGCAATTTTGGCAAGGGAGACCTGGAGAAATTTTGG 2280
Qy 2281 GACTGCTCTATGTTGAAAACATTTGCTTCTTCCAACTGCACTCCTTAAACTGC 2340
Db 2281 GACTGCTCTATGTTGAAAACATTTGCTTCTTCCAACTGCACTCCTTAAACTGC 2340
Qy 2341 CCTGTGGCTTTCTTGTCTCTCTCTTAAATAAACCTTACATTTATCAAGTCTCGAAGTA 2400
Db 2341 CCTGTGGCTTTCTTGTCTCTCTCTTAAATAAACCTTACATTTATCAAGTCTCGAAGTA 2400
Qy 2401 ATTAAGTTTATCTTCTGTTGGTAGTCCCACTTCTGCTGATGCTCAATCCCTCTCTAC 2460
Db 2401 ATTAAGTTTATCTTCTGTTGGTAGTCCCACTTCTGCTGATGCTCAATCCCTCTCTAC 2460
Qy 2461 ATCTTGTTCATCTCTCACTTTAGAGAGACTGCTGGTGGCTGAGAAAGCAAACTTACGTC 2520
Db 2461 ATCTTGTTCATCTCTCACTTTAGAGAGACTGCTGGTGGCTGAGAAAGCAAACTTACGTC 2520
Qy 2521 TGGACAAGATCAAAACACCCCAAGCTTGAATGCTCAATTAATCTCTGATGATGTCGAAAAACAG 2580
Db 2521 TGGACAAGATCAAAACACCCCAAGCTTGAATGCTCAATTAATCTCTGATGATGTCGAAAAACAG 2580


```
RESULT 4
US-09-495-050A-220
; Sequence 220, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED P
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 220
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 2603450CT1
US-09-495-050A-220

Query Match      26.1%; Score 710.4; DB 4; Length 723;
Best Local Similarity 99.7%; Pred. No. 1.5e-208;
Matches 722; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 1695 TTGGAGTGTGCACATAGCAGTCTCGACATCTTAAATGGGCTCATCGCAGCAG 1744
Db 1 TTGGAGTGTGCACATAGCAGTCTCGACATCTTAAATGGGCTCATCGCAGCAG 60

Qy 1745 TTTTCAGATCCCTCTGTACATTTCCCCCAATTAATGGGCTCATCGCAGCAG 1804
Db 61 TTTTCAGATCCCTCTGTACATTTCCCCCAATTAATGGGCTCATCGCAGCAG 120

Qy 1805 TGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGGCTGGTGTGATGCGTTCACTTTTG 1864
Db 121 TGAACATGCTCAGGGAGTCTCCAGTGCCTGCTGGCTGGTGTGATGCGTTCACTTTTG 180

Qy 1865 GCAGCTTTGCACGACATGCTGCTGGTGGGAGATGGGTTGGTCCATGCTATTGGTT 1924
Db 181 GCAGCTTTGCACGACATGCTGCTGGTGGGAGATGGGTTGGTCCATGCTATTGGTT 240

Qy 1925 TTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTCTCTTACTCTGCGACCCCTGGAGC 1984
Db 241 TTTTGTCCATTTTGTCTTCAGAAATCATCTGTTTCTCTCTTACTCTGCGACCCCTGGAGC 300

Qy 1985 GTGGGTTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCCTGA 2044
Db 301 GTGGGTTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTCTAGCCTGA 360

Qy 2045 AAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2104
Db 361 AAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

Qy 2105 GTGGCAGCAAGTATGGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2164
Db 421 GTGGCAGCAAGTATGGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480

Qy 2165 CCATGGGCTATAGTGTGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2224
Db 481 CCATGGGCTATAGTGTGCTCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540

Qy 2225 TTGCTACACCAAGCTCTACTGCAATTTGGACAGGAGACTGGAGAAATTTGGGACT 2284
Db 541 TTGCTACACCAAGCTCTACTGCAATTTGGACAGGAGACTGGAGAAATTTGGGACT 600

Qy 2285 GCTCTATGGTAAAAACACATGTCCTGTTGCTTTCACCAACTGCTATCCCTAACTGCCCC 2344
Db 1 GCTCTATGGTAAAAACACATGTCCTGTTGCTTTCACCAACTGCTATCCCTAACTGCCCC 444
```

```
Db 601 GCTCTATGGTAAAAACACATGTCCTGTTGCTTTCACCAACTGCTATCCCTAACTGCCCC 660
Qy 2345 TGGCTTTCTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2404
Db 661 TGGCTTTCTGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 719
Qy 2405 AGTT 2408
Db 720 AGTT 723

RESULT 5
US-09-799-451-723
; Sequence 723, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yuning
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: PC_FL_genes Version 2.0
; SEQ ID NO 723
; LENGTH: 2988
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2484)
US-09-799-451-723

Query Match      20.4%; Score 556.4; DB 4; Length 2988;
Best Local Similarity 56.8%; Pred. No. 1.8e-160;
Matches 1116; Conservative 0; Mismatches 821; Indels 27; Gaps 4;

Qy 562 TTGCAAGCCATGACCTTGGCCCTGACAAATAACACACATACCAGACTATGCTTTGGA 621
Db 145 TTGCAAGCCATGACCTTGGCCCTGACAAATAACACACATACCAGACTATGCTTTGGA 204

Qy 622 AACCTCTCCAGCTGGTGTGATGTTCTACATCTCCATTAAGTAATCCATCTCCCTGGGAAG 681
Db 205 GGTCTCTACAGCTGAAAAATCCTGATGCTGCAGAACCAATCAGCTGGGAGGAATCCCCGCA 264

Qy 682 AAATGCTTTGATGGCTCCACAGCTAGAGACTTTTAAATTTAAATTAACAACTTCAT 741
Db 265 GAGGCGCTGTGGGAGCTGCCGAGCTCGAGTCCGCTAGACCTGAATTTAAGAGCTGCAG 324

Qy 742 GAAATTCGCCCACTGCAATTTAGGACACTCTCCAACTCTTAAAGAACTAGGATTTTCATAGCAAC 801
Db 325 GAGTTCCTCTGGGCCATCCCGACCTTGGGAGACTGCGAGAACTGGGCTTCCATAACAAAC 384

Qy 802 AATATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAACCCCTTCTCTTATTACATACAT 861
Db 385 AACATCAGGTCGATACCTGAGAAAGCAATTTGTAGGCAACCCCTTCTCTTATTACATACAT 444
```



```
QY 862 TTCTATGACATCCCATCCAAATTTGTGGAGATCTGCTTTTCAACATTTTAACTGTAACCTA 921
Db 863 TTTCTATGACATCCCATCCAAATTTGTGGAGATCTGCTTTTCAACATTTTAACTGTAACCTA 921
QY 445 TTTTATGATAACCCATCCAGTTTGTGGAGATCGGCATTCAGTAGTACCTGCTAAATCTC 504
Db 446 TTTTATGATAACCCATCCAGTTTGTGGAGATCGGCATTCAGTAGTACCTGCTAAATCTC 504
QY 922 AGAACACTGACTCTGAATGCTGCTCACAAATAAATGAAATTTCCGTGATTTAACTGGAAT 981
Db 923 AGAACACTGACTCTGAATGCTGCTCACAAATAAATGAAATTTCCGTGATTTAACTGGAAT 981
QY 505 CACACACTATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
Db 506 CACACACTATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 564
QY 982 GCAAACTGAGAGTCTGACTTTTAACTGGAGACAGATCTCATCTCTCTCTCAAAACCGTTC 1041
Db 983 GCAAACTGAGAGTCTGACTTTTAACTGGAGACAGATCTCATCTCTCTCTCAAAACCGTTC 1041
QY 565 ACCAGCTGAGATCTGACCTGACCCGCGCAGGATCCGGGCTCCCATCGGGATG 624
Db 566 ACCAGCTGAGATCTGACCTGACCCGCGCAGGATCCGGGCTCCCATCGGGATG 624
QY 1042 TGAATCAGTTACTTAATCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
Db 1043 TGAATCAGTTACTTAATCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1101
QY 625 TGCCAAACAGTCCCGAGCTCGAGTCTCTGGAATCTCTCAAAATGAGGAGCTG 684
Db 626 TGCCAAACAGTCCCGAGCTCGAGTCTCTGGAATCTCTCAAAATGAGGAGCTG 684
QY 1102 CCCAGTTTTCAGTCTGCAAAAGCTTCAAGAAATTTGACCTTAAGCAATGAAATCTATC 1161
Db 1103 CCCAGTTTTCAGTCTGCAAAAGCTTCAAGAAATTTGACCTTAAGCAATGAAATCTATC 1161
QY 685 CCAGCTGCACAGGTGTGAGAAATTTGGAGAAATTCGGCTCCAAACACACCGCATCTGG 744
Db 686 CCAGCTGCACAGGTGTGAGAAATTTGGAGAAATTCGGCTCCAAACACACCGCATCTGG 744
QY 1162 GAAATTTAAAGTTGACACTTTCCAGCAGTTTGTCTTAGCTCCGATCGCTGAAATTTGGCTGG 1221
Db 1163 GAAATTTAAAGTTGACACTTTCCAGCAGTTTGTCTTAGCTCCGATCGCTGAAATTTGGCTGG 1221
QY 745 GAAATTTGAGCTGACACTTTCCAGCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGG 804
Db 746 GAAATTTGAGCTGACACTTTCCAGCAGCTGAGCTCCCTGCAAGCCCTGGATCTTAGCTGG 804
QY 1222 AACAATTTGCTATTATTCACCCCAATGCAATTTCCACTTTGCCATTCCTTAATAAAGCTG 1281
Db 1223 AACAATTTGCTATTATTCACCCCAATGCAATTTCCACTTTGCCATTCCTTAATAAAGCTG 1281
QY 805 AACGCCATCCGCTCCATCACCCTGAGGCTTCTCCACCCTGCACTCCCTGGTCAAGCTG 864
Db 806 AACGCCATCCGCTCCATCACCCTGAGGCTTCTCCACCCTGCACTCCCTGGTCAAGCTG 864
QY 1282 GACCTATGCTCAACCTCTGCTGCTTTTCTCTATTAACCTGGGTTACATGCTTAACTC 1341
Db 1283 GACCTATGCTCAACCTCTGCTGCTTTTCTCTATTAACCTGGGTTACATGCTTAACTC 1341
QY 865 GACCTGACAGACCAACAGCTGACCACTCCGCTGGCTGGCTGAGTCTGAGTAT 924
Db 866 GACCTGACAGACCAACAGCTGACCACTCCGCTGGCTGGCTGAGTCTGAGTAT 924
QY 1342 TTAATAATTAACAGGAAATCATGCTTTACAGAGCTTGATATCATCTGMAAACTTTCCAGAA 1401
Db 1343 TTAATAATTAACAGGAAATCATGCTTTACAGAGCTTGATATCATCTGMAAACTTTCCAGAA 1401
QY 925 CTGAGCTCAAGGNAACCTTCTCTCCAGGCTTCTCCAGGACAGATTTCCNAAA 984
Db 926 CTGAGCTCAAGGNAACCTTCTCTCCAGGCTTCTCCAGGACAGATTTCCNAAA 984
QY 1402 CTCAAGGTTATAGAATGCTTTATGCTTACAGTCTGCTGCAATTTGGAGTGTGAGAT 1461
Db 1403 CTCAAGGTTATAGAATGCTTTATGCTTACAGTCTGCTGCAATTTGGAGTGTGAGAT 1461
QY 985 CTGAGGATCTCGAGGTGCTTATGCTTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
Db 986 CTGAGGATCTCGAGGTGCTTATGCTTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1044
QY 1462 GCCTATGATTTCTTAATCAATGGAATAAGGTGACACAGCAGTATGACACCTTCAT 1521
Db 1463 GCCTATGATTTCTTAATCAATGGAATAAGGTGACACAGCAGTATGACACCTTCAT 1521
QY 1045 TTTCTTAAGGCTCTGCGCAGTGGGAGCTGAGGCTTCAAGCTTCACTTGAATGATGAGGAT 1104
Db 1046 TTTCTTAAGGCTCTGCGCAGTGGGAGCTGAGGCTTCAAGCTTCACTTGAATGATGAGGAT 1104
QY 1522 AAGAAA-----GATGCTGGAATGTTTCAAGCTCAAGATGAAGCTGACCTT 1566
Db 1523 AAGAAA-----GATGCTGGAATGTTTCAAGCTCAAGATGAAGCTGACCTT 1566
QY 1105 TCAAAAAGGCCCTGGGCTCTTGGCAGACAGCAGAGAACCACTATGACACGAGCTG 1164
Db 1106 TCAAAAAGGCCCTGGGCTCTTGGCAGACAGCAGAGAACCACTATGACACGAGCTG 1164
QY 1567 GAAGATTTCTGCTTGAATTTGAGGAGACCTGAAAGCCCTTCAATTCAGTGCAGTGTTC 1626
Db 1568 GAAGATTTCTGCTTGAATTTGAGGAGACCTGAAAGCCCTTCAATTCAGTGCAGTGTTC 1626
QY 1165 GATGAGCTCCAGCTGAGAT---GGAGGACTCAAGCCACACCCAGTGTCCAGTGTAGC 1221
Db 1166 GATGAGCTCCAGCTGAGAT---GGAGGACTCAAGCCACACCCAGTGTCCAGTGTAGC 1221
QY 1627 CTTTCCCGAGGCCCTTCAACCCCTGTGAAACCTGCTGAGCTGTGAGTGTGCTGATCAGAT 1686
Db 1628 CTTTCCCGAGGCCCTTCAACCCCTGTGAAACCTGCTGAGCTGTGAGTGTGCTGATCAGAT 1686
QY 1222 CTTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTTGAAGCTGGGGCATCCGCTG 1281
Db 1223 CTTACTCCAGGCCCTTCAAGCCCTGTGAGTACCTCTTTTGAAGCTGGGGCATCCGCTG 1281
QY 1687 GGAGTGTGACCATAGCAGTCTGCGCAGCTTACTTGTGATGCTTTGCTGACTTCAACAGTT 1746
Db 1688 GGAGTGTGACCATAGCAGTCTGCGCAGCTTACTTGTGATGCTTTGCTGACTTCAACAGTT 1746
QY 1282 GCGGTGTGGGCCATCGTGTGCTCTCGGCTCTGCAATGGACTGTGTGCTGTGACCGTG 1341
Db 1283 GCGGTGTGGGCCATCGTGTGCTCTCGGCTCTGCAATGGACTGTGTGCTGTGACCGTG 1341
QY 1747 TTC---AGATCCCTCTGTGACATTTCCCCCATTTAACTGTTAAATTTGGGCTCATCGCAGCA 1803
Db 1748 TTC---AGATCCCTCTGTGACATTTCCCCCATTTAACTGTTAAATTTGGGCTCATCGCAGCA 1803
QY 1342 TTTGCTGGGGGCTGTCCCCCTGCCCCGGTCAAGTTTGTGTTAGTGTAGTGTGAGTTCAGGC 1401
Db 1343 TTTGCTGGGGGCTGTCCCCCTGCCCCGGTCAAGTTTGTGTTAGTGTAGTGTGAGTTCAGGC 1401
QY 1804 GTGAACATGCTCAGCGGAGTCTCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863
Db 1805 GTGAACATGCTCAGCGGAGTCTCAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863
QY 1402 GCAACACCTTGAATGGCAATTTCTGCTGCTGCTTCTAGCTCAGTGTGATGCTGCTGCTT 1461
Db 1403 GCAACACCTTGAATGGCAATTTCTGCTGCTGCTTCTAGCTCAGTGTGATGCTGCTGCTGCTT 1461
QY 1864 GCGAGCTTGTGACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1923
Db 1865 GCGAGCTTGTGACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1923
QY 1462 GGTGAGTCTCTGAGTACGAGGCCCTGCTGGGAGACGGGGCTAGGCTGCCGGGCCACTGGC 1521
Db 1463 GGTGAGTCTCTGAGTACGAGGCCCTGCTGGGAGACGGGGCTAGGCTGCCGGGCCACTGGC 1521
```

```
QY 1924 TTTTGTCCATTTTGTCTTCAAGATCATCTGTTTCTGCTTACTCTGCGAGCCCTGGAG 1983
Db 1925 TTTTGTCCATTTTGTCTTCAAGATCATCTGTTTCTGCTTACTCTGCGAGCCCTGGAG 1983
QY 1522 TTTCTGGCAGTACTTTGGGTGCGAGCATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1581
Db 1523 TTTCTGGCAGTACTTTGGGTGCGAGCATCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1581
QY 1984 CGTGGTCTCTGCTGCAAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTG 2043
Db 1985 CGTGGTCTCTGCTGCAAAATATTTCTGCGGGCTATGGAAGTCCCTCTCTGCGGAGCTG 2043
QY 1582 TGCAGGCTCTCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1641
Db 1583 TGCAGGCTCTCGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1641
QY 2044 AAGTAATCATTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2103
Db 2045 AAGTAATCATTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2103
QY 1642 CGAGCAGGGTCTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1701
Db 1643 CGAGCAGGGTCTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1701
QY 2104 GGTGCGAGAGTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2157
Db 2105 GGTGCGAGAGTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2157
QY 1702 TCACTGGGAGATACGGGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1761
Db 1703 TCACTGGGAGATACGGGGCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1761
QY 2158 CCAGCACCATGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2217
Db 2159 CCAGCACCATGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2217
QY 1762 CCAGCAGCTTGGGCTTCCCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1821
Db 1763 CCAGCAGCTTGGGCTTCCCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1821
QY 2218 ATGACATTTGCTACACCAAGCTTACTGCAATTTTGGACAGGAGACCTGGAGAAATTT 2277
Db 2219 ATGACATTTGCTACACCAAGCTTACTGCAATTTTGGACAGGAGACCTGGAGAAATTT 2277
QY 1822 GTGGCGGTGCTTACATCAAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1881
Db 1823 GTGGCGGTGCTTACATCAAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1881
QY 2278 TGGGACTGCTTATGTTAAACACATTTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2337
Db 2279 TGGGACTGCTTATGTTAAACACATTTGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2337
QY 1882 TGGGACTGCGCATGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1941
Db 1883 TGGGACTGCGCATGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1941
QY 2338 TGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2397
Db 2339 TGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2397
QY 1942 TGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2001
Db 1943 TGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2001
QY 2398 GTAATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2457
Db 2399 GTAATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2457
QY 2002 GCGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2061
Db 2003 GCGCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2061
QY 2458 TACATCTTGTTCATCTTCACTTTTAAAGAGAGATCTGGTGGAGCT 2501
Db 2459 TACATCTTGTTCATCTTCACTTTTAAAGAGAGATCTGGTGGAGCT 2501
QY 2062 TACCTGCTCTTCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2105
Db 2063 TACCTGCTCTTCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2105
```

RESULT 6

US-09-495-050A-214
Sequence 214, Application US/09495050A
Patent No. 6492505

GENERAL INFORMATION:
APPLICANT: Roopa, Reddy

APPLICANT: Guegler, Karl, J.

APPLICANT: Au-Young, Janice

TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED

FILE REFERENCE: PA-0013 US

CURRENT APPLICATION NUMBER: US/09/495,050A

CURRENT FILING DATE: 2000-01-31

PRIOR APPLICATION NUMBER: 60/118,318

PRIOR FILING DATE: February 1, 1999

NUMBER OF SEQ ID NOS: 305

SOFTWARE: PERL Program

SEQ ID NO 214

LENGTH: 2612

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. 6492505 2470285CT1

US-09-495-050A-214

Query Match 19.8%; Score 540; DB 4; Length 2612;
Best Local Similarity 55.1%; Pred. No. 1.9e-155;
Matches 1121; Conservative 0; Mismatches 885; Indels 30; Gaps 2;

QY 468 CTGTTTTCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 527
Db 112 CAGTTTTCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 171

QY 528 AATCCCGCTCAGGCTTTTAGAGTTTATCGGCAATGCAAGCCATGACCTTGGCCCTGAA 587
Db 172 GGTGCTGTGCACCCCTCAGCAATCTGCCACCCCTACAGGCGTGACCCCTGGCTCTCAA 231
QY 588 CAAATAACACACATACCAAGACTATGCTTTGGAAACCTCTCCAGCTTGGTAGTTCTACA 647
Db 232 CAAGATCTCAAGCATCCCTGACTTTTGGCAATTTACCAACCTTTCAAGCCCTGGTAGTTCTGCA 291
QY 648 TCTCCATAAATAAGAAATCCACTCCCTGGGAAAGAAATGCTTTGATGGCTTCCACAGCCT 707
Db 292 TCTTCATAAATAAATAAGAGCTGAGTCAACACTGTTTGGTGACTAGATAAGCCT 351
QY 708 AGAGACTTTAGATTTAAATTAACATAACTTGTAGTAATTTCCCACTGCAATAGGACACT 767
Db 352 GGAGACCTTAGACTTGAATTAATAACTTTGGGGAATTTCCCTCAGGCTATTAAGCCCT 411
QY 768 CTCCAACCTTAAAGAACTAGGATTTATAGCAACAATAATCAGGTCGATACCTGAGAAAGC 827
Db 412 TCCTAGCCTTAAAGAGCTAGGATTTATAGTAATTTCTATTTCTGTATCCCTGATGGAGC 471
QY 828 ATTTGTAGGCAACCTTCTCTTATTAACAATACATTTCTATGACAAATCCCAATTTGT 887
Db 472 ATTTGATGGTAATCCACTCTTAAGAACTATACATTTGTATGATAATCCTCTGTCTTTGT 531
QY 888 TGGGAGATCTGTTTCAACATTTTACCTGAACCTGAAACACTGACTCTGAATGGTGCCTC 947
Db 532 GGGGAACCTCAGCATTTCAAAATTTATCTGATCTTCATTTCCCTAGTCAATCTGGTGGCAAG 591
QY 948 ACAATAACTGAATTTCTGATTTAACTTGAACCTGCAACCTGGAGAGCTCTGACTTTAAAC 1007
Db 592 CATGGTGACAGTTCCTCCCAATCTTACAGGAACCTGTCACCTGGAAAGCTCTGACTTTGAC 651
QY 1008 TGGAGCAGATCTCATCTTCTTCCCTCAACCGTCTGCAATCAGTTACCTAATCTCCAAGT 1067
Db 652 AGGTACAAGATTAAGCAGCATACCTTAATAATTTGTGCCAAGAAACAAAGATGCTTAGGAC 711
QY 1068 GCTAGATCTGCTTACACCTTATAGAGATTTTACCAGATTTTTCAGTCTGCCAAAGCT 1127
Db 712 TTTGAGCTTGTCTTACAATAATATAAGAGACCTTCCAAAGTTTAAATGGTTGCCATGCTCT 771
QY 1128 TCAGAAAATTCACCTTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCCAGCA 1187
Db 772 GGAAGAAATTTCTTACAGCGTAACTCAATCTACCAATTAAGGAAGCACTTTCAAGG 831
QY 1188 GTTGCTTAGCCTCCGATCTCGTGAATTTGGCTTGGAAACAAATATGCTATTTTCAACCCAA 1247
Db 832 CCTGATATCTCTAAGGATCTAGATCTGAGTAGAAACCTGATACATGAAATTCACAGTAG 891
QY 1248 TGCAATTTTCCACTTTGGCATCCCTAATAAGCTGGACCTATCGTCCACCTCTGTCGTC 1307
Db 892 AGCTTTTGGCCACACTTGGGCCAAATAAATCACTAGATGTAAGTTTCAATGAATTAACCTC 951
QY 1308 TTTTCTATAACTGGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCCTT 1367
Db 952 CTTTCTTAGGAGGCTGAATGGCTTAATCACTGAAATCTGTGGGCAACTCAAGCT 1011
QY 1368 ACAGAGCTTGATATCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGC 1427
Db 1012 GAAAGAGCCTTAGCAGCAAAAAGACTTTTGTAAACCTCAGGTCTTTATCAGTACCATATGC 1071
QY 1428 TTACAGGTCTGCTGATTTGGAGTGTGTGAGNATGCGCTATAGATTTTCTAATCAATGAA 1487
Db 1072 TTATCAGTGTGTGATTTTGGGGTTGTG-----ACTCTTATGCAAAATTTAA 1119
QY 1488 TAAAGTGACAAACAGCAGTATGGAGCAGCTTTCATAAGAAAGATGCTGGAATGTTTCAGGC 1547
Db 1120 CACAGAGATTAACAGCTTCAGGACCAC-----AGTGGGACAGGA 1161
QY 1548 TCAAGATGAAAGTGAACCTTGAAGATTTCTGCTTGAAGTCTTGGAGGAACCTGGAAGCCCT 1607
Db 1162 GAAAGGTACTGCTGATGAGCAAAATGTCAAGACACTCTTGAAAATGAAGACATAGTCA 1221

QY 1608 TCATTCAGTGCAGTGTTTCACTTCCCGAGGCCCTTCAAAACCTGTGTGAACACCTGCTTGA 1667
Db 1222 AATTAATTTATCATTTGTACACCTTCAACAGGTGCTTTTAAAGCCCTGTGNATATTACTGGG 1281
QY 1668 TGGCTGGCTGATCAGAAATTTGGAGTGTGGACCATAGCAGTTCTGGCACCTTATGTAATGC 1727
Db 1282 AAGCTGGATGATCTGCTTACTGTGTGTTTCAATTTCTTGGTGTGCAATTTATTTTCAACCT 1341
QY 1728 TTTGGTGACTTCAACAGTTTTTCAGATCCCTCTCTGTACATTTTCCCCCATTAACATGTTAAT 1787
Db 1342 GCTTGTATTTTAAACAACATTTGCACTCTGTACATCACTGCTTTCGTCCAAATTTGTTAT 1401
QY 1788 TGGGCTCATCGCAGCAGTGAACATGCTCACGGGAGTCTCAGTGCCTGCTGGCTGGTGT 1847
Db 1402 AGGCTTGATTTCTGTCTTAACCTTATTCATGGGAATCTATCTGGCATCTCTAACTTTCT 1461
QY 1848 GGATGCGTTCACTTTTGGCAGCTTTTGCAACATGCTGCTGCTGGTGGGAGAAATGGGTTGG 1907
Db 1462 TGATGCTGTGCTGGGGCAGATTCGCTGAATTTGGCAATTTGGTGGGAAACCTGCGAGTGG 1521
QY 1908 TTGCCATGTCAATGGTGTGTTTGTCCATTTTGTCTTCAAGATCATCTGTTTCTGCTTAC 1967
Db 1522 CTGCAAAAGTAGTGGGTTTCTTGCAGTTTTCTCTCAGAAAGTGCCATATTTTATTAAT 1581
QY 1968 TCTGGCAGCCCTGAGCGTGGGTTCTCTGTGAAATATTTCTCAAAATTTGAAAACGAAAGC 2027
Db 1582 GCTAGCAACTGTCGAAAGAGCTTATCTGCAAAAGATATATGAAATTTGGAGAGCAA 1641
QY 2028 TCCATTTTCTAGCTGAAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTTGAACAATGGC 2087
Db 1642 TCATCTCAACAGTTCTGGGTTGCTGCTGCTTTCCTAGGTGCTACAGTAGCAGG 1701
QY 2088 CGGATTTCCCTGCTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTTGCC 2147
Db 1702 CTGTTTTCCCTTTTCCATAGAGGGGAATATTCGCAATCACCCCTTTGTTGCCATTTCC 1761
QY 2148 TTTTGGGAGCCAGCAGCACCATGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTG 2207
Db 1762 TACAGGTGAAACGCCATCATTAGGATTCACCTGTAAAGTTAGTGTCTATTAACCTACTAGC 1821
QY 2208 CTTCTCATGATGACCAATTCCTTACACCAAGCTCTACTGCAATTTTGGCAAGGAGACCT 2267
Db 1822 ATTTTATTAATGGCCGTATCTACATAAGCTATATACTGCAACTTGGAAAAAGAGGACCT 1881
QY 2268 GGAGAAATTTGGGAGCTCTCTATGTAAGCAATTTGCTGCTGCTGCTTCCACCACTG 2327
Db 1882 CTCAGAAACTCAACAATCTAGCATGATTAAGCATGCTGGCTTAATCTTCAACAATG 1941
QY 2328 CATCTAAACTGCTGCTGCTTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2387
Db 1942 CATCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2001
QY 2388 CAGTCTGAAGTAATTAAGTTTATCTTCTGCTGGTGGTAGTCCCACTTCTGCTGATGTCTCAA 2447
Db 2002 CAGCCCGAAATAATGAAGTCTGTACTCTGATATTTTTCATTTGCTGCTGCTGCTGCTGCTG 2061
QY 2448 TCCCTCTCTCTACATCTGTTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2503
Db 2062 TCCAGTCTCTGATGTTTCTTCAACCCCAAGTTTAAAGAGACTGGAAGTTACTGA 2117

RESULT 7

US-08-866-757-1
; Sequence 1, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA

STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/866,757
FILING DATE: 30-MAY-1997

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-70055
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4203 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-866-757-1

Query Match 18.8%; Score 511.2; DB 2; Length 4203;

Best Local Similarity 54.9%; Pred. No. 2.2e-146;

Matches 1090; Conservative 0; Mismatches 863; Indels 31; Gaps 3;

QY 520 TTAACAGAAATCCCGTCAGGCTTTTAAAGTTTATGGCATTCGCAAGCCATGACCTTG 579
DB 56 TTGACGGAGGTGCGTGTGACCCCTCAGCAATCTGCCACCTACAGGCGCTGACCTG 115
QY 580 GCCCTGAACAAATAACACACATACAGACTATGCTTTGGAACTCTCAGCTTGGTA 639
DB 116 GCTCTCAACAAGATCTCAAGCATCCCTGACTTTGCAATTTACCACTTTCAAGCCCTGTA 175
QY 640 GTTCTACATCTCCATAACAATAGATCCATCCCTCGGGAAGAAATGCTTTCATGGGCTC 699
DB 176 GTTCTGCATCTTCAACAATAAATAAGAGCTGAGTCAACACTGTTTGTATGGACTA 235
QY 700 CACAGCTAGAGACTTTAGATTAAATTAACAATAACCTTGTATGAATTCGCCACTGCAATT 759
DB 236 GATAACCTGGAGACCTTAGACTTGAATTATAATACTTTGGGGAAATTTCTCAGGCTATT 295
QY 760 AGGACACTCTCAACTTAAGAACTAGATTTCATAGCAACAATATCAGGTTCGATACCT 819
DB 296 AAAGCCCTTCTAGCCTTAAAGAGCTAGATTTCATAGTAATTCATTTCTGTTATCCCT 355
QY 820 GGAAGAAGCAATTTGAGGCAACCTTCTCTTATTAACAATATCTATGATCAATCCCATC 879
DB 356 -ATGAGAGCAATTTGATGGTAATCCACTTTAAGAACTATACATTTGTATGTAATCCCTG 414
QY 880 CAATTTGTTGGAGATCTGCTTTTCAACATTTACCTGAACTAAGAACTGACTCTGAAT 939
DB 415 TCTTTTGGGGAATCTAGCACTTCAATTTATCTGATCTTCTTCTTCTTCTTCTTCTTCT 474
QY 940 GGTGCTTCAAAATAACTGAATTTCTGATTTTAACCTGGAACCTGCAAACTCGAGAGCTG 999
DB 475 GGTGCAAGCATGGTCAGCAGTTCCCAATCTTACAGGAATGTCCCACTGGAAGTCTG 534
QY 1000 ACTTTAACTGGAGCAGATCTATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1059
DB 535 ACTTTGACAGGTACAAAGATAGCAGCATACCTTAATAATTTTGTGTCAAGAAACAAAGATG 594

QY 1060 CTCCAAGTGTAGATCTCTCTTACAACTATTAGAAGATTTACCAGTTTTTTCAGTCTGC 1119
DB 595 CTTAGGACTTTGGACTTGTCTTACAAATAATAAGAGACCTTCCAAGTTTTTAATGTTGC 654
QY 1120 CAAAGCTTCAAGAAATGACCTAAGACATATGAATCTAGAAATTAAGTTGACACT 1179
DB 655 CATGCTCGGAAGAAATTTCTTTACAGCGTAATCAAACTTACCAATAAAGGAAGCAACC 714
QY 1180 TTCCAGCAGTTGCTTAGCTCCGATCGCTGAATTTGGCTTGGAAACAAATTCCTATTATT 1239
DB 715 TTTCAAGCCCTGATATCTCTAAGGATTTAGATCTGAGTAGAAGAACTGATACATGAAT 774
QY 1240 CACCCCAATGATTTTCCACTTTGCAATCCCTAATAAAGCTGAGCACTATGTCACACCTC 1299
DB 775 CACAGTAGAGCTTTTGCCACACTTGGGCAATAAACAACCTAGATGTAAGTTTCAATGAA 834
QY 1300 CTGTGCTCTTTTCTTCTAATAAATCTGGGTATACATGGTTTAACTCACTTAAATTAACAGAAAT 1359
DB 835 TTAACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 894
QY 1360 CATGCCCTTACAGAGCTTGTATATCATCTGAAACTTTCCAGAACTCAAGCTTATAGAAATG 1419
DB 895 TTTCAAGCTGAAGAGCCCTTAGCAGCAAGACCTTTGTTAACTCAGCTCTTTATCAGTA 954
QY 1420 CCTTATGCTTACAGTGTCTGTGCAATTTGGAGTGTGTGAGATGCTTAAGATTTCTAAT 1479
DB 955 CCATATGCTTATCAGTGTCTGTGCAATTTTGGGGTGTG-----ACTTTATGCA 1002
QY 1480 CAATGGAATAAAGGTGACACAGCAGATGATGACGACCTTCAAGAAAGATGCTGGAATG 1539
DB 1003 AATTTAAACACAGAAAGATAACAGCCCTCCAGGACCAACAGT-----GTG 1044
QY 1540 TTTCAAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTTGTGCTTTGAGGAAGACCTG 1599
DB 1045 GCACAGGAAGAGTACTGCTGTATGCAAGCAATGTCAAGCACTCTTGAAGATGAGAA 1104
QY 1600 AAAGCCCTTCAATCAGTGCAGTGTTCACCTTTCCCGAGGCCCTTCAAAACCTGTGAACAC 1659
DB 1105 CATAGTCAATAATAATTTATCCATTGTACACTTTCAACAGGTGCTTTTAAAGCCCTGTGAAT 1164
QY 1660 CTGCTTGTATGCTGGCTGATCAGAAATGGAGTGTGACCATAGCAGTCTTGGCATTACT 1719
DB 1165 TTACTGGGAAGCTGGATGATTCGTCTTACTGTGTGTTTCAATTTCTTGGTTGCAATTTT 1224
QY 1720 TGTATGCTTTTGTGACTTCAACAGTGTTCAGATCCCTCTGTACATTTCCCCCATTTAA 1779
DB 1225 TTCAACCTGCTTGTATTTTAAACAATTTGCACTTGTATCATCTGCTTCTGTCGCAAA 1284
QY 1780 CTGTTAATTTGGGTCATCGCAGCAGTGAACATGCTCAGGGAGTCTCCAGTGCCTGCTG 1839
DB 1285 TTGTTTATAGCTTGAATTTCTGTGTCTAACTTATTCATGGGAATCTATACTGGCATCTTA 1344
QY 1840 GCTGGTGTGATGGCTTCACTTTTGGGAGCTTTTGAACATGCTGCTGCTGCTGGGAGAT 1899
DB 1345 ACTTTTCTTGTATGCTGCTGCTGGGCAAGATTCGCTGAATTTGGCAATTTGGTGGGAACT 1404
QY 1900 GGGGTTGGTTCGCAATGCTTTTGTGTCATTTTGTGTCATGCAATCATCTGCTTTTC 1959
DB 1405 GGCAGTGTCTCAAGTAATCTGGGTTTCTGCAAGTTTCTCCTCAGAAAGTGCCTATTT 1464
QY 1960 CTGCTTACTCTGGCAGCCCTGGAGCGTGGGTTCTCTGTGAATATTTCTGCAAAATTTTGA 2019
DB 1465 TTATTAAATGCTAGCAACTGTGCAAGAAAGCTTATCTGCAAAAGATATAATGAAGATGG 1524
QY 2020 ACGAAAGCTTCAATTTTCTAGCCTGAAGATATCAATTTTGTCTGTGCTGCTGCTGCTG 2079
DB 1525 AAGAGCAATCATCTCAAAACAGTTCGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1584
QY 2080 ACCATGCGCAGTTCCTGCTGGGTGGCAGAGTATGCGGCTCCCTCTCTGCTGCTG 2139
DB 1585 GTAAAGAGCTGTTTTCCTCTTTTCCATAGAGGGGAATATCTGCAATCACCCCTTTGTTG 1644


```
Qy 2140 CTTTGGCTTTTGGGAGCCAGACCATGAGCTACATGGTCTCATCTTCTGCTCAAT 2199
Db |||||
Qy 1645 CCAITTCCTCAGGTGAACACGCCATCAATTAGGATTCAGTAAAGCTTAGTCTATTAAC 1704
Db |||||
Qy 2200 TCCCTTTGCTTCCCTCATGATGACCAATTCGCTTACACCAAGCTCTACTGCAATTTGGACAAG 2259
Db |||||
Qy 1705 TCACTAGCAATTTTATTAATGCGCTTATCTACACTAAGCTATATCTGCAACTTGGAAAAA 1764
Db |||||
Qy 2260 GGAGACCTGGAGAAATATTTGGGACTGCTCTATGTAAGAAACACATTTGCCCTGTTGCTCTTC 2319
Db |||||
Qy 1765 GAGGACCTCTCAGAAAACTCAAACTAGCAATTAAGCAATGCTGCTTGGCTTAATCTTC 1824
Db |||||
Qy 2320 ACCAACTGCATCTTAACTGCGCTGCTGCTTCTTCTGCTCTCTCTCTTAAATAAACCTT 2379
Db |||||
Qy 1825 ACCAAATGCATCTTTTCTGCGCTGCTGCGCTTTTCTCAATTTGCAACATGATCACTGCA 1884
Db |||||
Qy 2380 ACATTTATCAGTCTCTGAAGTAAATTAAGTTTATCTCTGCTGGTGGTGGTCCCACTTCCCTGCA 2439
Db |||||
Qy 1885 ATCTCTATCAGCCCGGAAATTAAGTCTGTTACTCTGATATTTTTCATTTGCTGCTGCT 1944
Db |||||
Qy 2440 TGTCTCAATCCCTTCTCTCATCTTGTCTCAATCTCTCACTTTAAGGAGGATCTGGTGAGC 2499
Db |||||
Qy 1945 TGCCTGAATCCAGTCTGATGTTTCTTCAACCAAAAGTTTAAAGAAGACTGGAAGTTA 2004
Db |||||
Qy 2500 CTGA 2503
Db |||||
Qy 2005 CTGA 2008
Db |||||

RESULT 8
US-09-153-593-1
; Sequence 1, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGMA, DEREK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153.593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4203
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-153-593-1

Query Match 18.8%; Score 511.2; DB 3; Length 4203;
Best Local Similarity 54.9%; Pred. No. 2.2e-146;
Matches 1090; Conservative 0; Mismatches 863; Indels 31; Gaps 3;

Qy 520 TTAACAGAAATCCCGTCCAGGCTTTTAGAGTTTATCGGCAATTCGCAAGCCATGACCTTG 579
Db |||||
Qy 56 TTGACGGAGGTGCTGTGTCACCCCTCAGCAATCTGCCACCCCTACAGGCGCTGACCCGTG 115
Db |||||
Qy 580 GCCCTGAACAAATATACACCATACAGACTATGCTTTGGAAACCTCTCCAGCTTTGGTA 639
Db |||||
Qy 116 GCTCTCAACAAGATCTCAAGCATCTGCTGACTTTGGATTTACCAACCTTTCAAGCTGGTA 175
Db |||||
Qy 640 GTTCTACATCTCCATAACAATAGAAATCCATCCCTGGGAAAGAAATGTTTGTATGGGCTC 699
Db |||||
Qy 176 GTTCTGCACTTTCATAACAATAAAATTAAGAGCCTGAGTCAACACTGTTTGTATGGACTA 235
Db |||||
Qy 700 CAGACCTTAGACCTTTAGATTTAAATTAACATTAACCTTTGATGATTTCCCACTGCAATT 759
Db |||||
Qy 236 GATAACCTGGAGACCTTAGACTTGAATTTAATAAATTTGGGGAAATTTCCCTCAGGCTATT 295
Db |||||
Qy 760 AGGACACTCTCAACCTTAAAGAACTAGGATTTTCATAGCAACAATATCAGGTCGATACCT 819
Db |||||
```

```
Db 296 AAAGCCCTTCTAGCCCTTAAAGAGCTAGGATTTTCATAGTAATTCATTTCTGTGTTATCCCT 355
Qy 820 GAGAAAGCATTTGTAGCAACCCCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATC 879
Db |||||
Qy 356 -ATGAGCAATTTGATGTAATCCACTCTTAAAGAACTATACATTTGTATGATTAATCCCTGTG 414
Db |||||
Qy 880 CAATTTGCTGGGAGATCTGCTTTTCAACATTTACTCTGAACTTAAGAAACACAGCTCTGAAT 939
Db |||||
Qy 415 TCTTTTGTGGGAACTCAGCAATTTCAAAATTTATCTGATCTTCAATCCCTAGTCAATTCGT 474
Qy 940 GGTGCTCTCACAAATAACTGAATTTCTGATTTAACTGGAACCTGCAACCTGGAGAGCTG 999
Db |||||
Qy 475 GGTGCAAGCATGCTGAGCAGTTCCTCCAACTTACAGGAACTGCTCCACCTGGAAAGCTG 534
Db |||||
Qy 1000 ACTTTAACTGGAGACAGATCTCATCTCTCTCAACCCGCTCTGCAATCAGTTACCTAAT 1059
Db |||||
Qy 535 ACTTTGACAGTACAAGATAAGCAGCATACCTAATAATTTGTGTCAAGAACAAAAAGATG 594
Db |||||
Qy 1060 CTCCAAGTGTAGATCTGCTTACCAACTATTAGAAGATTTACCAGTCTTTTTCAGTCTGC 1119
Db |||||
Qy 595 CTTAGGACTTTGGACTTGTCTTCAATAATAAAGAGACCTTCCAGTCTTTTAATGGTTC 654
Qy 1120 CAAAAGCTTCAGAAAAATGACCTAAGACATAAATAAGAAATCTACGAAATTTAAAGTTGACACT 1179
Db |||||
Qy 655 CATGCTCTGGAAGAAATTTCTTTACAGCGTAATCAAAATCTACCAATAAAGGAAGGCACC 714
Db |||||
Qy 1180 TTTCCAGCAGTGTGTTAGCTCCGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATT 1239
Db |||||
Qy 715 TTTCAAGGCTGATATCTCTAAGGATTTAGATCTGAGTAGAAACCTGTATACATGAATTT 774
Db |||||
Qy 1240 CACCCCAATGCATTTTCCACTTTGCCATCTCTAATAAGCTGGACCTATGCTCCAACTC 1299
Db |||||
Qy 775 CACAGTAGACTTTTGGCCACTTGGGCCAATAACTAACTAGATGTAGTTTCAATGAA 834
Qy 1300 CTGTGCTTTTCTTAACTGAGTGTGAGTGTGAGAACTTTCCAGAACTCAAGTTTATAGAATG 1359
Db |||||
Qy 835 TTAACCTTCTTCTTCCAGGAGGCTGAAATGGGCTAAATCAACTGAACCTGGTGGGCAAC 894
Db |||||
Qy 1360 CATGCTTTACAGACTTGATATCATCTGAAACCTTTCCAGAACTCAAGTTTATAGAATG 1419
Db |||||
Qy 895 TTCAAGCTGAAGAAGCCTTTAGCAGCAAAAGACTTTGTTAACTCAGGTCTTTATCAGTA 954
Qy 1420 CTTTATGCTTACCAAGTCTGTGCAATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAAT 1479
Db |||||
Qy 955 CCATATGCTTATCAGTCTGTGCAATTTGGGGTGTG-----ACTCTATGCA 1002
Db |||||
Qy 1480 CAATGGAATAAGGTGACAAACAGCAGTATGGAACCTTTCATAGAAGAAAGTGTGGAATG 1539
Db |||||
Qy 1003 AATTTAAACACAGAAAGATAACAGCCTCCAGACCACAGT-----GTG 1044
Db |||||
Qy 1540 TTTCAGGCTCAAGATGAACCTGACCTTGAAGATTTCTGCTTGTGACTTTGAGGAAGACCTG 1599
Db |||||
Qy 1045 GCACAGAGAAGAGGTACTGCTGATGCAAGCAAAATGTCACAGCACTCTTGAATAAGAA 1104
Qy 1600 AAAGCCCTTCAATTCAGTGCAGTGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAAACAC 1659
Db |||||
Qy 1105 CATAGTCAATAATATTCATTTGACACTTCAACAGTGTCTTTAAGCCCTGTGAATAT 1164
Qy 1660 CTGCTTGTGCTGCTGATCAGAAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTACT 1719
Db |||||
Qy 1165 TTAATGGAAGCTGATGATTCGTCTTACTGTGTGTTTCAATTTCTTGGTGTGCTATTATTT 1224
Qy 1720 TGTAACTGTTTGTGACTTCAAGATTTTCAGATCCCTCTGTATACATTTCCCCCATTA 1779
Db |||||
Qy 1225 TTCAACCTGCTGTTTATTTTAAACAACATTTGCACTTTGTATCATCACTGCTTCTGCAAA 1284
Qy 1780 CTGTTAATTTGGGTCATCGCAGCAGTGAACATGCTCAAGGAGTCTCAGTGCCTGCTG 1839
Db |||||
Qy 1285 TTGTTTATAGCTTGAATTTCTGTCTTAACTTATTCATGGAACTATATCTGATCCTTA 1344
Qy 1840 GCTGGTGGAGTGGTTTCACTTTTGGAGCTTTTGCACGACATGCTGCTGGTGGAGAAAT 1899
Db |||||
Qy 1345 ACTTTTCTGATGCTGTCTCTGGGGCAGATTCGCTGAATTTTGGCAATTTGGTGGGAACT 1404
Db |||||
```


QY 1900 GGGGTTGGTGGCCATGTCATGGTGTGTTTGTGTCATATTTTGGTTCAGAAATCATCTGTTTC 1959
DB 1405 GGCAGTGGCTGCAAAAGTAACCTGGGTTTCTTGAGTTTCTCTCCAGAAAGTGCCATATTT 1464
QY 1960 CTGCTTACTCTGGCAGCCCTGGAGCGTGGTTCTCTGTGAATAATTTCTGCAAAATTTGAA 2019
DB 1465 TTATTAATGCTAGCAACTGTGCAAGAGCTTATCTGCAAAAGATATTAATGAAATGGG 1524
QY 2020 ACGAAGCTCCATTTTCTAGCCCTGAAAGTAATCATTTTGTCTGTGCGCTGTGCGCCTTG 2079
DB 1525 AAGAGCAATCATCTCAACAGTTCCGGTGTGCGCCCTTTTGGCTTTCTAGGTGCTACA 1584
QY 2080 ACCATGGCGGCGATTCCTCCCTGCTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCGCTG 2139
DB 1585 GTAACAGGCTGTGTTTCCCTTTTCCATAGAGGGGAATATCTGCAATCAACCCCTTTGTTTG 1644
QY 2140 CCTTTGGCTTTTGGGAGCCAGCAGCCATGGCTACATGGTCTCATCTTGTCTCAAT 2199
DB 1645 CCAATTTCTACAGGTGAACGCCATCATTTAGGATTCACGTAACTTAGTGTATTAAC 1704
QY 2200 TCCCTTTGCTTCTCATGATGACCACTTGCCTACACCAAGCTCTACTGCAATTTGGACAAG 2259
DB 1705 TCACCTAGCATTTTATTAATGCGGTTATCTACACTAAGCTATAGTCACTTGAAGAAA 1764
QY 2260 GGAGACCTGGAGAATATTTGGGACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTC 2319
DB 1765 GAGGACCTCTCAGAAAACCTCAACAATAGCATGATTAAGCATGCTGCTTGGCTAATCTTC 1824
QY 2320 ACCAATGCATCTAACTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2379
DB 1825 ACCAATGCATCTTTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1884
QY 2380 ACAATTTATCAGTCTGAGTAATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2439
DB 1885 ATCTATCAGCCCGCAATATAGAGCTGTACTCTGATATTTTCCATTTGCTGCTGCT 1944
QY 2440 TGCTCAATCCCTTCTCTACATCTGTTCAATCTCTCAATCTCTCAATCTCTCAATCTCTCAATCT 2499
DB 1945 TGCTCAATCCAGTCTCTGATGTTTCTCAACCCCAAGTTTAAAGAAAGACTGGAAGTTA 2004
QY 2500 CTGA 2503
DB 2005 CTGA 2008

RESULT 9
; Sequence 75, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030PI
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 75
; LENGTH: 2703
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-482-273-75

Query Match 4.6%; Score 125.8; DB 4; Length 2703;

Best Local Similarity 59.3%; Pred. No. 2.1e-27;
Matches 214; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 2143 TTGCCCTTTTGGGAGCCCGCAGCACCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTC 2202
DB 10 TTTCCTACAGGTGAAACGCCCATATTAGGATTCACGTAAACGTTAGTGTCTATTAAACTCA 69
QY 2203 CTTTGTCTTCTCATGATGACCAATTTGCCCTACACCAAGCTCTACTGCAATTTGGACAAGGA 2262
DB 70 CTAGCATTTTATTAATGGCGTTATCTACACTAAGCTATATCTGCAACTTATGCAAAAGAG 129
QY 2263 GACCTGGAGAATATTTGGGAGCTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTCC 2322
DB 130 GACCTCTCAGAAAACCTCAAACTTAGCATGATTAAGCATGTCGCTTGGCTAATCTTCACC 189
QY 2323 AACTGCATCTTAAACTGCGCTGCTGCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2382
DB 190 AATTGCATCTTTTCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 249
QY 2383 TTTATCAGTCTCCTGAAGTAATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2442
DB 250 TCTATCAGCCCGCAATTAATGAAGTCTGTTACTCTGATATTTTTCATTTGCTGCTGCTGCTGCT 309
QY 2443 CTCATCCCTTCTCTACATCTTGTTCATCTCTCACTTTAAAGGAGGATCTGCTGAGCCTG 2502
DB 310 CTGAATCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 369
QY 2503 A 2503
DB 370 A 370

RESULT 10
US-08-866-757-3
; Sequence 3, Application US/08866757
; Patent No. 5858716
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMA, DERK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/866,757
; FILING DATE: 30-MAY-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70055
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 473 base pairs
; TYPE: nucleic acid


```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-866-757-3

Query Match      4.5%; Score 123.8; DB 2; Length 473;
Best Local Similarity 61.2%; Pred. No. 2.2e-27;
Matches 200; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

Qy 2168 TGGGCTACATGGTCGCTCTCATCTTGTCTCAATTCCTCTTGGTCTCTCATGATGACCAATG 2227
Db 20 TAGGATTCACCTGAACGTTAGTGCTATTAAACTCACTAGCATTTTATTAAATGGCCGTTA 79

Qy 2228 CCTACACCAAGCTCTACTGCAATTTGGACAGGAGGACCTGGAGATATTTGGGACTGCT 2287
Db 80 TCTACACTAAGCTATACTGCAACTTTGGAAAAAGAGGACCTCTCAGAAAACTCACAATCTA 139

Qy 2288 CTATGGTAAACACATTCGCCCTGTGTCTTCCACCAACTGCATCTAACTGACCTGAGG 2347
Db 140 GCATGATTAAAGATGCTGCTGGCTTAATCTTCAACCAATTCATCTATCAGCCCGGAAATATGAAGT 199

Qy 2348 CTTTCTTCTCCTCTCTCTCTTTTAAATAAACCTTACATTTTATCAGTCTCTGAAATTAAGT 2407
Db 200 CGTTTTTTTTCATTTGACCAATTCATCTGCAATCTCTATCAGCCCGGAAATATGAAGT 259

Qy 2408 TTATCCCTCTGGTGTAGTCCCACTTCTCTGATCTCTCAATCCCTCTCTTACATCTTGT 2467
Db 260 CTGTCTACTCTGATATTTTTTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319

Qy 2468 TCAATCCCTCACTTTAAAGGAGGATCTGG 2494
Db 320 TCAACCCAAAGTTTAAAGAGGACTGG 346

RESULT 12
US-09-482-273-101
; Sequence 101, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; CURRENT FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 101
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-482-273-101

Query Match      4.2%; Score 115.4; DB 4; Length 2709;
Best Local Similarity 59.3%; Pred. No. 3.4e-24;
Matches 214; Conservative 0; Mismatches 146; Indels 1; Gaps 1;

Qy 2143 TTGCTTTTGGGGAGCCAGCACCATGCGGTACATGGGTCTCTCATCTTCTCAATTC 2202
Db 10 TTCTCTACAGGTGAACGCCCATCATATTAGGATTCATCTGTAACGTTAGTCTATTAACTCA 69

Qy 2203 CTTTGTCTTCTCATGATGACCATTTGCTTACACCAAGCTCTACTGCAATTTTGGACAAAGGA 2262
Db 70 CTAGCATTTTATTAAATGGCGGTTATCTACACTAAGCTATATCTGCACTTGGAAAAAGAG 129

Qy 2263 GACCTGGAGATAATTTGGGACTGCTCTATGTGTAACACATTTGCCCTTCTTCTTACC 2322
Db 130 GACCTCTCAGAAACTCACATCTAGCATGATTAAAGCATGTCGCTTGGCTAATCTTACC 189

Qy 2323 AACTGCATCTTAAACTGCGCTTCTTGTGCTTCTCTCTCTTTTAAATTAACCTTACA 2382
Db 190 AATTGCATCTTTTCTGCGCTGTCGCTTTTTCATTTGCAACCATTCATGATCCTGCAATC 249

Qy 2383 TTTATCAGTCTGAGTAATTAAGTTTATCTTCTGTTGGTAGTCCCACTTCTTCTGATGT 2442
Db 250 TCTATCAGCCCGGAAATAAGTCTGTTACTCTGATATTTTTTCCATGCTGCTGCTGCTG 308

Qy 2443 CTCAATCCCTTCTCTACATCTTGTTCATCTCTCACTTTTAAAGGAGGATCTGGTGAAGCTG 2502
Db 309 CTGAATCCAGTCTCTGATGTTTCTTCAACCCAAAGTTTAAAGAGGACTGGAAGTTACTG 368

Qy 2503 A 2503
Db 369 A 369

Query Match      4.5%; Score 123.8; DB 3; Length 473;
Best Local Similarity 61.2%; Pred. No. 2.2e-27;
Matches 200; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

Qy 2168 TGGGCTACATGGTCGCTCTCATCTTGTCTCAATTCCTCTTGGTCTCTCATGATGACCAATG 2227
Db 20 TAGGATTCACCTGAACGTTAGTGCTATTAAACTCACTAGCATTTTATTAAATGGCCGTTA 79

Qy 2228 CCTACACCAAGCTCTACTGCAATTTGGACAGGAGGACCTGGAGATATTTGGGACTGCT 2287
Db 80 TCTACACTAAGCTATACTGCAACTTTGGAAAAAGAGGACCTCTCAGAAAACTCACAATCTA 139

Qy 2288 CTATGGTAAACACATTCGCCCTGTGTCTTCCACCAACTGCATCTAACTGACCTGAGG 2347
Db 140 GCATGATTAAAGATGCTGCTGGCTTAATCTTCAACCAATTCATCTATCAGCCCGGAAATATGAAGT 199

Qy 2348 CTTTCTTCTCCTCTCTCTCTTTTAAATAAACCTTACATTTTATCAGTCTCTGAAATTAAGT 2407
Db 200 CGTTTTTTTTCATTTGACCAATTCATCTGCAATCTCTATCAGCCCGGAAATATGAAGT 259

Qy 2408 TTATCCCTCTGGTGTAGTCCCACTTCTCTGATCTCTCAATCCCTCTCTTACATCTTGT 2467
Db 260 CTGTCTACTCTGATATTTTTTCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319

Qy 2468 TCAATCCCTCACTTTAAAGGAGGATCTGG 2494
Db 320 TCAACCCAAAGTTTAAAGAGGACTGG 346

RESULT 11
US-09-153-593-3
; Sequence 3, Application US/09153593A
; Patent No. 6174994
; GENERAL INFORMATION:
; APPLICANT: ELSHOURBAGY, NABIL A
; APPLICANT: LI, XIAOTONG
; APPLICANT: BERGSMAN, DEBK J
; TITLE OF INVENTION: NOVEL 7TM RECEPTOR (H2CAA71)
; FILE REFERENCE: GH-70055-1
; CURRENT APPLICATION NUMBER: US/09/153,593A
; CURRENT FILING DATE: 1998-09-15
; EARLIER APPLICATION NUMBER: 08/866,757
; EARLIER FILING DATE: 1997-05-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 473
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (397) (400) (432) (459)
US-09-153-593-3

Query Match      4.5%; Score 123.8; DB 3; Length 473;
Best Local Similarity 61.2%; Pred. No. 2.2e-27;
Matches 200; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

Qy 2168 TGGGCTACATGGTCGCTCTCATCTTGTCTCAATTCCTCTTGGTCTCTCATGATGACCAATG 2227
Db 20 TAGGATTCACCTGAACGTTAGTGCTATTAAACTCACTAGCATTTTATTAAATGGCCGTTA 79

Qy 2228 CCTACACCAAGCTCTACTGCAATTTGGACAGGAGGACCTGGAGATATTTGGGACTGCT 2287
Db 80 TCTACACTAAGCTATACTGCAACTTTGGAAAAAGAGGACCTCTCAGAAAACTCACAATCTA 139

Qy 2288 CTATGGTAAACACATTCGCCCTGTGTCTTCCACCAACTGCATCTAACTGACCTGAGG 2347
```


RESULT 13

US-09-016-434-1469
; Sequence 1469, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1469:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2413 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9903759
; US-09-016-434-1469

Query Match 3.9%; Score 106.8; DB 4; Length 2413;
Best Local Similarity 45.0%; Pred. No. 1.4e-21;
Matches 399; Conservative 0; Mismatches 487; Indels 0; Gaps 0;
QY 1616 TCGAGTGTTCACCTCCCGAGCCCTTCAAAACCTGTGAACACCTGTTGATGGCTGGC 1675
DB 1257 TGGTGTGACCCCAAGTCCGATGAGTTCAACCCGTGTGAACACATAATGGGCTACAAGT 1316
QY 1676 TCGACAGAAATGGAGTGTGGACATAGCAGTTCGSCACTTACTTGTAAATGCTTTGGTGA 1735
DB 1317 TCGTGAAGATGGTGTGGTTCGTTGATGCTGCTCTCTGGGCAATGCTTTTGCC 1376
QY 1736 CTTCAACAGTTTTTCAGATCCCTCTGTGTACATTTCCCGCCATTAACCTGTTAATGGGGTCA 1795
DB 1377 TCGTTATTCTCTCACCAGCCACTACAAGTGAACCTCCCGCTTTCTCATGTGCAACC 1436
QY 1796 TCGCAGCAGTGAACATGCTCAGGGAGTCTCAGTGCCGTGCTGGCTGTGTGGATGGCT 1855
DB 1437 TGGCCCTTTCCGATTTCTGCATGGGAGTGTACCTGCTCTCATCGCCTCTGTAGACCTCT 1496
QY 1856 TCACCTTTGGCAGCTTTGACACACATGGTCCCTGTGGGAGAAATGGGGTTGGTGGCCATG 1915
DB 1497 ACACCTACTCTGAGTACTACAACCATGCCATCGACTGGCAGACAGGCCCTGGGGTCAACA 1556

QY 1916 TCATTTGGTTTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTGCTTACTCTGGCAG 1975
DB 1557 CGGCTGGTTTTCTTCACTGTCTTTTGAAGCGAGTTATCGGTGTATACGTGACGGTCATCA 1616
QY 1976 CCCTGGAGCGTGGGTCTCTGTGAAATATTCTGCAAAATTTGAAACGAAAGCTCCATTTT 2035
DB 1617 CCCTGGAGCGCTGGTATGCGCATCACTTGGCATGGCCCTGACCGGAAGATCCGCTCA 1676
QY 2036 CTAGCCTGAAAGTAATCAATTTTGTCTCTGTGCCCTCTGCGCTTGACCATGGCCGAGTTTC 2095
DB 1677 GGCACGCATGTGCCATCATGGTTGGGGCTGGGTGTGCTCTCTTCTCGCCCTGCTTC 1736
QY 2096 CCCTGCTGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCGCTGCTTGGCTTTTGGGG 2155
DB 1737 CTTTGGTGGGAATAAGTAGCTATGCCAAAGTCAGTATCTGCGCTGCCCATGGACACCGAGA 1796
QY 2156 AGCCCGACGACCATGGGCTACATGGTGGCTCTCATCTGTCTCAATTTCCCTTTGCTTCTCA 2215
DB 1797 CCCCTCTGCTGGCATATATTGTTTGTCTGACGCTCAACATAGTTGCTTGGCTTGGTCA 1856
QY 2216 TGATGACCAATTGCTTACCAAGCTCTACTGCAATTTGGACAAGGGAGACCTGGGAATA 2275
DB 1857 TCGTCTGCTGCTGTTATGTGAAGATCTACATCACAGTCCGAATCCGACACACCCAGGGG 1916
QY 2276 TTTGGGACTGCTCTATGTGTAACACATGCGCTGCTGCTTCTTACCAACTGCTCACTAA 2335
DB 1917 ACAAGATACCAAAATTTGCCAAGGATGGGTGTGATCTTCCCGACTTCCAGCTGCA 1976
QY 2336 ACTGCCCTGTGGCTTTCTTGTCTCTCTCTTTAATAAACCTTACATTTATCAGTCCCTG 2395
DB 1977 TGGCCCCAATCTCATTTCTATGCTGTGTCAGCAATTTCTGAACAGGCTCTCATCATCTG 2036
QY 2396 AAGTAATTAAGTTTATCCTTCTGCTGGTGTAGTCCCACTTCTGTCATGTCTCAATCCCTTC 2455
DB 2037 GCAACTCCAAATCTTGTGCTGCTTCTTCTATCAATTAACCTCTGTGCAATCCATTC 2096
QY 2456 TCTACATCTTGTTCATCTCCTCACTTTAAGGAGGATCTGTTAGGCTT 2501
DB 2097 TCTATGCTATTTTCCAAAGGCTTCCAGAGGGATGTGTTTCATCCT 2142

RESULT 14

US-07-741-453A-62
; Sequence 62, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/741,453A
; FILING DATE: 19911015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 3710 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-07-741-453A-62

Query Match 3.7%; Score 100.2; DB 3; Length 3710;
Best Local Similarity 45.3%; Pred. No. 2.2e-19;
Matches 403; Conservative 0; Mismatches 483; Indels 3; Gaps 1;

QY 1616 TGCAGTGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAACACCTGCTGTGATGGCTGGC 1675
Db |||||
QY 1519 TGGTGTGACCCCNAGTCCGATGAGTTCAACCCGTGTGAACACATATATGGCTACAAGT 1578
Db |||||
QY 1676 TGATCAGAAATGGAGTGTGGACCAATGAGTGTCTGGCACTTACTTGTAAATGCTTTGGTGA 1735
Db |||||
QY 1579 TCCTGAGAAATGGTGTGGTTCGTTAGTCTGTGGCTCTCTGGGCAATGCTTTGTGTC 1638
Db |||||
QY 1736 CTTCAACAGTTTTAGATCCCTCTGTATATTTCCCCATTTAAACTGTAATTTGGGTCA 1795
Db |||||
QY 1639 TGCCTTATCTCTCCACGACCACTACAAACTGAACGTCCTCCCGCTTTCTCATGTGCAACC 1698
Db |||||
QY 1796 TCGCAGCAGTGAACATGCTCAGCGAGTCTCCAGTGGCTGTGGCTGGTGTGATGGT 1855
Db |||||
QY 1699 TGGCTTTGGGATTTCTGCATGGGATGTACTGCTCTCATGCCCTCTGTAGACCTCT 1758
Db |||||
QY 1856 TCACCTTTTGGCAGCTTTGACGACATGTTGCTGGTGGGAGAAATGGGTTGGTTGCCATG 1915
Db |||||
QY 1759 ACACCTCACTCTGAGTACTACAAACCATGCCAGTGGCAGACAGGCCCTGGGTGCAACA 1818
Db |||||
QY 1916 TCATTTGGTTTTTGTCCATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGGCAG 1975
Db |||||
QY 1819 CGGCTGGTTTCTTCACTGCTTTGCAAGCGAGTTATCGGTGTATACGCTGACGGTCA 1878
Db |||||
QY 1976 CCCTGGAGCTGGGTTCTCTGTGAATATCTGCAAAATTTGAAACGAAAGCTCCATTTT 2035
Db |||||
QY 1879 CCCTGGAGCGTGTATGCGATCCATCCTTCGCCATGCGCTGGACCGGAAGATCGCCTCA 1938
Db |||||
QY 2036 CTAGCCTGAAAGTAAATATTTTGTCTGTGCCCTGCTGCCCTTGACCATGCGCGAGTTC 2095
Db |||||
QY 1939 GGCACGATGTGCCATCATGTTGGGGGCTGGGTTTGTCTTCTCTCGCTCGCCCTGCTTC 1998
Db |||||
QY 2096 CCCTGCTGGTGGCAGCAAGTATGCGGCTCCCTCTCTGCTGCTGCTTTGGGG 2155
Db |||||
QY 1999 CTTTGGTGGGAATAAGTAGTACTGCAAAAGTCAGTATCTGCTGCCATGACACCGAGA 2058
Db |||||
QY 2156 AGCCAGCAGCATGGCTACATGTTGCTGCTCATCTGCTCAATTTCCCTTTGCTTCCTCA 2215
Db |||||
QY 2059 CCCCTTCTGTGCAATATATGTTTGTCTGACGCTCAAATAGTTGCTTCTGCTCA 2118
Db |||||
QY 2216 TGATGACCAATGCTCTACCAAGCTCTACTGCAATTTGGACAA---GGGAGACCTGGAGA 2272
Db |||||
QY 2119 TCGTCTGCTGTTATGTAAGATCTATACATCACAGTCCGAAATCCGAGTACAAACCCAG 2178
Db |||||
QY 2273 ATATTTGGGACTGCTATATGTAACACATGTCCTGCTGCTTTCACCAACTGCATCC 2332
Db |||||
QY 2179 GGGACAAAGATACCAAAATGCGAAGAGATGGCTGTGTTGATCTTACCGCACTTCATAT 2238
Db |||||
QY 2333 TAAACTGCCCTGTGGCTTTCTGCTCTCTCTCTTAAATAAAGCTTACATTTATCAGTC 2392
Db |||||
QY 2239 GCATGGCCCCAATCTCATTTCTATGCTGTGAGCAATTTCTGAAAGGCTCTCATCACTG 2298
Db |||||
QY 2393 CTGAAGTAATTAAGTTTATCTCTTCTGGTGGTAGTCCCACTTCTCTGTCATGTCTCAATCCC 2452
Db |||||
QY 2299 TTAGCAACTCCAAAATCTTGTGGTACTTCTTATCCACTTAACTCTCTGTGCAATCCAT 2358
Db |||||

QY 2453 TTCTCTATATCTTGTTCAAATCTCTCACCTTAAAGGAGGATCTGTTGAGCCT 2501
Db |||||
QY 2359 TCCTCTATGCTATTTTCCACCAAGCCCTTCCAGAGGATGTGTTTATCTCT 2407
Db |||||

RESULT 15
US-07-741-453A-57
; Sequence 57, Application US/07741453A
; Patent No. 6228597
; GENERAL INFORMATION:
; APPLICANT: PARMENTIER, MARC
; APPLICANT: LIBERT, FREDERIC
; APPLICANT: DUMONT, JACQUES
; APPLICANT: VASSART, GILBERT
; TITLE OF INVENTION: POLYPEPTIDES HAVING THYROTROPIN-RECEPTOR
; TITLE OF INVENTION: ACTIVITY, NUCLEIC ACID SEQUENCES CODING FOR SUCH RECEPTORS
; TITLE OF INVENTION: AND POLYPEPTIDES, AND APPLICATIONS OF THESE POLYPEPTIDES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN DARBY & CUSHMAN
; STREET: 1615 L STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: U.S.A.
; ZIP: 20036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07741,453A
; FILING DATE: 19911015
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16773
; REFERENCE/DOCKET NUMBER: 91913/1107/US/ST
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 861-3000
; TELEFAX: (202) 822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4417 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-07-741-453A-57

Query Match 3.7%; Score 100; DB 3; Length 4417;
Best Local Similarity 45.1%; Pred. No. 2.8e-19;
Matches 413; Conservative 0; Mismatches 500; Indels 3; Gaps 1;

QY 1616 TGCAGTGTTCACCTTCCCGAGGCCCTTCAAAACCTGTGAACACCTGCTGTGATGGCTGGC 1675
Db |||||
QY 1249 TGGTGTGACTCTTAAGTCAGATGAGTCAACCTCTGAAGACATATATGGCTACAAGT 1308
Db |||||
QY 1676 TGATCAGAAATGGAGTGTGGACCAATGAGTGTCTGGCACTTACTTGTAAATGCTTTGGTGA 1735
Db |||||
QY 1309 TCCTGAGGATTTGGTGTGGTGTGTTAGTCTGCTGGCTCTCTGGGCAATGCTTTGTGTC 1368
Db |||||
QY 1736 CTTCAACAGTTTTAGATCCCTCTGTATATTTCCCCATTTAAACTGTTAATTTGGGTCA 1795
Db |||||
QY 1369 TGATCGTCTCTTACCAGTCACTACAAATTGACTGTGCCACGCTTTCTCATGTGCAACT 1428
Db |||||
QY 1796 TCGCAGCAGTGAACATGCTCAGCGAGTCTCCAGTGGCTGTGGCTGGTGTGATGGT 1855
Db |||||
QY 1429 TGGCTTTGAGATTTCTGCATGGGATGTATCTGCTCTCATGCCCTCCGCTAGACCTCT 1488
Db |||||
QY 1856 TCACCTTTGGCAGCTTTTGACGACATGTTGCTGGTGGGAGAAATGGGTTGTTGCCANG 1915
Db |||||
QY 1489 ACACCTCATCTGAGTACTTACAAACCATGCCATCGACTGGCAGACAGAGGCCCTGGGTGAACA 1548
Db |||||

This Page Blank (uspto)

1	2724	100.0	2724	15	US-10-251-385-263	Sequence 263, App
2	2724	100.0	2724	15	US-10-325-567A-431	Sequence 421, App
3	2724	100.0	2724	17	US-10-174-456-A	Sequence 4, Appl
4	2724	100.0	2724	20	US-10-751-736-21	Sequence 21, Appl
5	2724	100.0	2880	17	US-10-595-027-111A	Sequence 111A, App
6	2724	100.0	2880	21	US-10-482-029-157	Sequence 157, App
7	2724	100.0	2880	21	US-10-451-237-48	Sequence 48, Appl

Db 1 ATGGACACTCCCGGCTCGGTGCTGCTGCTGCTTGGCTGTGCTGCTGCTGCTGCGGACC 60
Qy 61 GGGGCACTCTCCAGGCTCTGGTGTGTTGTGTGAGGGGCTGCCCCACACACTGTCATTGC 120
Db 61 GGGGCACTCTCCAGGCTCTGGTGTGTTGTGTGAGGGGCTGCCCCACACACTGTCATTGC 120
Qy 121 GAGCCCGAGCGAGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Db 121 GAGCCCGAGCGAGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGGAGCTG 180
Qy 181 CTTTCCAACTCAGGCTCTTCACTCTTACCTAGACCTCAGTATGAACAAACATCAGTCAG 240
Db 181 CTTTCCAACTCAGGCTCTTCACTCTTACCTAGACCTCAGTATGAACAAACATCAGTCAG 240
Qy 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTGAGAGGATTAAGTCTTTGGGGAAAC 300
Db 241 CTGCTCCCGAATCCCTGCCAGTCTCCGCTTCTGAGAGGATTAAGTCTTTGGGGAAAC 300
Qy 301 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCCTTACAGTCTTAAAGTCTTTATG 360
Db 301 GCTCTGACATACATTTCCCAAGGAGCATTTCACTGGCCCTTACAGTCTTAAAGTCTTTATG 360
Qy 361 CTGCAGAAATATCAGCTAAGACACGTACCCACAGAACTCTGCAGAAATTTGCGAAGCCTT 420
Db 361 CTGCAGAAATATCAGCTAAGACACGTACCCACAGAACTCTGCAGAAATTTGCGAAGCCTT 420
Qy 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 421 CAATCCCTGCGTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Qy 481 CTGCATTCCTTGAGGACCTGCTGGCTGGATGCAATGCTTTAAACAGAAATCCCGTCCAG 540
Db 481 CTGCATTCCTTGAGGACCTGCTGGCTGGATGCAATGCTTTAAACAGAAATCCCGTCCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAAACAAATACACCAC 600
Db 541 GCTTTTGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAAACAAATACACCAC 600
Qy 601 ATACAGACTATGCTTTGGAAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAAT 660
Db 601 ATACAGACTATGCTTTGGAAAACCTCTCCAGCTTGGTAGTTCTACATCTCCATAACAAAT 660
Qy 661 AGAATCCACTCCCTGGGAAGAAATGCTTTCATGSGCTCCACAGCCTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAGAAATGCTTTCATGSGCTCCACAGCCTAGAGACTTTAGAT 720
Qy 721 TTAAATTAACAATAACCTTGTGAAATTCGCCACTGCAATTAGGACACTCTCCAACTTAA 780
Db 721 TTAAATTAACAATAACCTTGTGAAATTCGCCACTGCAATTAGGACACTCTCCAACTTAA 780
Qy 781 GAACTAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Qy 841 CTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCATCCATTTGTTGGAGATCTGCT 900
Db 841 CTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCATCCATTTGTTGGAGATCTGCT 900
Qy 901 TTTCAACAATTTACCTGAACCTAAGAACACTGACTCTGAAATGCTGCTCAAAATACGAA 960
Db 901 TTTCAACAATTTACCTGAACCTAAGAACACTGACTCTGAAATGCTGCTCAAAATACGAA 960
Qy 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGTCTGACTTTAACTGGAGCAAGATC 1020
Db 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGTCTGACTTTAACTGGAGCAAGATC 1020
Qy 1021 TCATCTCTTCTCAAAACCTGCGAATCAGTTTACCTTAATCTTCAAGTCTGATCTGCT 1080
Db 1021 TCATCTCTTCTCAAAACCTGCGAATCAGTTTACCTTAATCTTCAAGTCTGATCTGCT 1080
Qy 1081 TACAACTTATAGAAAGTTTACCCAGTCTTTCAGTCTGCCAAAGCTTTCAGAAAATTGAC 1140
Db 1081 TACAACTTATAGAAAGTTTACCCAGTCTTTCAGTCTGCCAAAGCTTTCAGAAAATTGAC 1140

Qy 1141 CTAAGACATAATGAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
Db 1141 CTAAGACATAATGAATCTACGAAATTTAAAGTTGACACTTTCCAGCAGTTGCTTAGCCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTTGGAAACAAAATTTGCTATTATTACACCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTTGGAAACAAAATTTGCTATTATTACACCCCAATGCAATTTTCCACT 1260
Qy 1261 TTGCCATCCCTAAATAAAGCTGGACCTATCGTCCAACTCTCTGTGCTTTTCTTATTAAT 1320
Db 1261 TTGCCATCCCTAAATAAAGCTGGACCTATCGTCCAACTCTCTGTGCTTTTCTTATTAAT 1320
Qy 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
Db 1321 GGGTTACATGGTTTAACTCACTTAAATTAACAGGAATCATGCTTTACAGAGCTTGATA 1380
Qy 1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGTGT 1440
Db 1381 TCATCTGAAAACCTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGTGT 1440
Qy 1441 GCATTTGGAGTGTGTGAGAAATGCCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAA 1500
Db 1441 GCATTTGGAGTGTGTGAGAAATGCCCTATAAGATTTCTAATCAATGGAATAAAGGTGACAA 1500
Qy 1501 AGCAGTATGACGACCTTCAATAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Db 1501 AGCAGTATGACGACCTTCAATAGAAAGATGCTGGAATGTTTCAAGGCTCAAGATGAACGT 1560
Qy 1561 GACCTTCAAGATTTCTGCTTGACTTTGAGAAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1561 GACCTTCAAGATTTCTGCTTGACTTTGAGAAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Qy 1621 TGTTCACCTTCCCAAGCCCTTCAAAACCTGTGAAACACCTGCTTGTATGGCTGCTGATC 1680
Db 1621 TGTTCACCTTCCCAAGCCCTTCAAAACCTGTGAAACACCTGCTTGTATGGCTGCTGATC 1680
Qy 1681 AGAATTCGAGTGTGGACCATAGCAGTCTGSCACTTACTTGTATGCTTTGCTGACTTCA 1740
Db 1681 AGAATTCGAGTGTGGACCATAGCAGTCTGSCACTTACTTGTATGCTTTGCTGACTTCA 1740
Qy 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAACTGTTAAATTTGGGGTTCATCGCA 1800
Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTTAACTGTTAAATTTGGGGTTCATCGCA 1800
Qy 1801 GCAGTGAACTGCTCAGGGAGTCTCCAGTCCCGTGTGCTGCTGCTGCTGCTGCTTCACT 1860
Db 1801 GCAGTGAACTGCTCAGGGAGTCTCCAGTCCCGTGTGCTGCTGCTGCTGCTGCTTCACT 1860
Qy 1861 TTTGGCAGCTTTGTCAGCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 TTTGGCAGCTTTGTCAGCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy 1921 GGTTTTTTGTCCATTTTGTCTTCAAGAAATCACTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTGTCTTCAAGAAATCACTGTTTTTCTGCTTACTCTGGCAGCCCTG 1980
Qy 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGTCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTTCCATTTTCTAGC 2040
Qy 2041 CTGAAAGTAAATCAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2041 CTGAAAGTAAATCAATTTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Qy 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGGCC 2160
Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGGCC 2160
Qy 2161 AGCAACATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTCCTTCTCATGATG 2220
Db 2161 AGCAACATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTTCCTTCTCATGATG 2220

QY 2221 ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATATTGG 2280

Db 2221 ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTGGAGAAATATTGG 2280

QY 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGTCTTTCCACCACTGCATCCCTAAACTGC 2340

Db 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGTCTTTCCACCACTGCATCCCTAAACTGC 2340

QY 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2400

Db 2341 CCTGTGGCTTTCTTGTCTCTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2400

QY 2401 ATTTAAGTTTATCTTCTGTGGTGGTAGTCCACTTCTCTGCAATGTCATATGCCCTTCTCTAC 2460

Db 2401 ATTTAAGTTTATCTTCTGTGGTGGTAGTCCACTTCTCTGCAATGTCATATGCCCTTCTCTAC 2460

QY 2461 ATCTTGTTCATCTCTCACTTTTAAAGAGAGATCTGTGTGAGCCCTGAGAAAGCAAACTACGTC 2520

Db 2461 ATCTTGTTCATCTCTCACTTTTAAAGAGAGATCTGTGTGAGCCCTGAGAAAGCAAACTACGTC 2520

QY 2521 TGGCAAGATCAAACCAACCCCAAGCTTGTATGTCAATTAACCTCTGATGATGTCGAAAAACAG 2580

Db 2521 TGGCAAGATCAAACCAACCCCAAGCTTGTATGTCAATTAACCTCTGATGATGTCGAAAAACAG 2580

QY 2581 TCCTGTGACTCAACTCAAGCTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTG 2640

Db 2581 TCCTGTGACTCAACTCAAGCTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTG 2640

QY 2641 CCTCCAGTTCCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700

Db 2641 CCTCCAGTTCCGTGCCATCACCAGCTTATCCAGTGACTGAGAGCTGCCATCTTCTCTCT 2700

QY 2701 GTGGCATTTGTCCCATGTCTCTAA 2724

Db 2701 GTGGCATTTGTCCCATGTCTCTAA 2724

RESULT 2

US-10-225-567A-421

; Sequence 421, Application US/102255567A

; Publication No. US20030113798A1

; GENERAL INFORMATION:

; APPLICANT: LifeSpan Biosciences

; APPLICANT: Brown, Joseph P.

; APPLICANT: Roush, Christine L.

; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS

; FILE REFERENCE: 1920-4-4

; CURRENT APPLICATION NUMBER: US/10/225,567A

; CURRENT FILING DATE: 2001-12-19

; PRIOR APPLICATION NUMBER: 60/257,144

; PRIOR FILING DATE: 2000-12-19

; NUMBER OF SEQ ID NOS: 2392

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 421

; LENGTH: 2724

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-225-567A-421

Query Match 100.0%; Score 2724; DB 15; Length 2724;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGACACCTCCGGCTCGGTGCTCTCTGTCTTGGCTGTGCTGCTGAGCTGGCGACC 60

Db 1 ATGAGACACCTCCGGCTCGGTGCTCTCTGTCTTGGCTGTGCTGCTGAGCTGGCGACC 60

QY 61 GGGGGGAGCTCTCCAGGTCTGGTGTGTCTGAGGGGCTGCCCCACACACTGTCATTTC 120

Db 61 GGGGGGAGCTCTCCAGGTCTGGTGTGTCTGAGGGGCTGCCCCACACACTGTCATTTC 120

QY 121 GAGCCCGACGGCAGGATGTTGCTCAGGGTGAGCTGCTCCGACCTGGGGCTCTCGGAGCTG 180

QY 61 GGGGACGCTCTCCAGGCTGTGGTGTGCTGAGGGGCTGCCACACACTGTCTATTGC 120
Db |||||
QY 109 GGGGACGCTCTCCAGGCTGTGGTGTGCTGAGGGGCTGCCACACACTGTCTATTGC 168
Db |||||
QY 121 GAGCCGACGGCAGGATGTGCTCAGGGTGGAGTGTCTCCGACCTGGGGCTCTCGAGCTG 180
Db |||||
QY 169 GAGCCGACGGCAGGATGTGCTCAGGGTGGAGTGTCTCCGACCTGGGGCTCTCGAGCTG 228
QY |||||
QY 181 CCTTCAACCTCAGGCTCTACCTCCCTACCTAGACCTCAGTATGAAACAATCAGTCAG 240
Db |||||
QY 229 CCTTCAACCTCAGGCTCTACCTCCCTACCTAGACCTCAGTATGAAACAATCAGTCAG 288
QY |||||
QY 241 CTGCTCCCGAATCCCTCCGAGTCTCCGCTTCTCGAGGAGTACGTCCTGGGGAAC 300
Db |||||
QY 289 CTGCTCCCGAATCCCTCCGAGTCTCCGCTTCTCGAGGAGTACGTCCTGGGGAAC 348
QY |||||
QY 301 GCTCTGACATACATTTCCCAAGGGAGCAATTCACCTGGCCCTTTACAGTCCTTAAAGTCTTTATG 360
Db |||||
QY 349 GCTCTGACATACATTTCCCAAGGGAGCAATTCACCTGGCCCTTTACAGTCCTTAAAGTCTTTATG 408
QY |||||
QY 361 CTGCGAATTAATCAGCTAAGACAGTACCCACAGAAAGCTCTGCAGAAATTTGGGAAGCCTT 420
Db |||||
QY 409 CTGCAGAAATAATCAGCTAAGACAGTACCCACAGAAAGCTCTGCAGAAATTTGGGAAGCCTT 468
QY |||||
QY 421 CAATCCCTGCGTCTGAGTGTAAACACACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGGC 480
Db |||||
QY 469 CAATCCCTGCGTCTGAGTGTAAACACACATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGGC 528
QY |||||
QY 481 CTGCATTCCTTGAGCACCTGTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGTCCAG 540
Db |||||
QY 529 CTGCATTCCTTGAGCACCTGTGGCTGGATGACAAATGCGTTTAAACAGAAATCCCGTCCAG 588
QY |||||
QY 541 GCTTTTGAAGTTTATCGGATTTGAAGCAATGCAAGCAGTACCTTGGCCCTGAAACAAATACACCAC 600
Db |||||
QY 589 GCTTTTGAAGTTTATCGGATTTGAAGCAATGCAAGCAGTACCTTGGCCCTGAAACAAATACACCAC 648
QY |||||
QY 601 ATACAGACTATGCTTTGGAAACCTCTCCAGCTGTGTAGTCTACATCTCCATAACAAT 660
Db |||||
QY 649 ATACAGACTATGCTTTGGAAACCTCTCCAGCTGTGTAGTCTACATCTCCATAACAAT 708
QY |||||
QY 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 720
Db |||||
QY 709 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTAGAGACTTTAGAT 768
QY |||||
QY 721 TTAATTTACAAATAACTTGTATGAATTTCCCACTGCAATTAGACACTCTCCAACTTAAA 780
Db |||||
QY 769 TTAATTTACAAATAACTTGTATGAATTTCCCACTGCAATTAGACACTCTCCAACTTAAA 828
QY |||||
QY 781 GAACTAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
Db |||||
QY 829 GAACTAGGATTTATAGCAACAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 888
QY |||||
QY 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 900
Db |||||
QY 889 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCCATCCCAATTTGTTGGAGATCTGCT 948
QY |||||
QY 901 TTTGAAACATTTACCTGAACTAAGAACACTGACTCTGAATGTGCTCTCAAAATAACTGAA 960
Db |||||
QY 949 TTTGAAACATTTACCTGAACTAAGAACACTGACTCTGAATGTGCTCTCAAAATAACTGAA 1008
QY |||||
QY 961 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Db |||||
QY 1009 TTTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1068
QY |||||
QY 1021 TCATCTCTTCTCAAACCTGTGCAATCAGTTTACCTAATCTCAAAGTCTAGATCTGTCT 1080
Db |||||
QY 1069 TCATCTCTTCTCAAACCTGTGCAATCAGTTTACCTAATCTCAAAGTCTAGATCTGTCT 1128
QY |||||
QY 1081 TACAACCTATTAGAGATTTACCCAGTTTTTCAGTCTGCGAAAGCTTTCAGAAATTCAC 1140
Db |||||
QY 1129 TACAACCTATTAGAGATTTACCCAGTTTTTCAGTCTGCGAAAGCTTTCAGAAATTCAC 1188
QY |||||
QY 1141 CTAAGACATTAATGAAATCTACGAAATTAAGTTGACACTTTTCCAGCAGTGTGCTTAGCCTC 1200

Db |||||
QY 1189 CTAAGACATTAATGAAATCTACGAAATTAAGTTGACACTTTTCCAGCAGTGTCTTAGCCTC 1248
QY |||||
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTAATTCACCCCAATGCAATTTTCACT 1260
Db |||||
QY 1249 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTAATTCACCCCAATGCAATTTTCACT 1308
QY |||||
QY 1261 TTGCGCATCCCTAAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATTAAT 1320
Db |||||
QY 1309 TTGCGCATCCCTAAATAAAGCTGGACCTATCGTCCAACTCTGCTGCTTTTCCCTATTAAT 1368
QY |||||
QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA 1380
Db |||||
QY 1369 GGGTTACATGTTTAACTCACTTAAATTAACAGGAAATCATGCTTTACAGAGCTTGATA 1428
QY |||||
QY 1381 TCATCTGAATAAATTTTCAGAACTCAAGGTTATAGAAATGCTTTTACAGTGTCTGT 1440
Db |||||
QY 1429 TCATCTGAATAAATTTTCAGAACTCAAGGTTATAGAAATGCTTTTACAGTGTCTGT 1488
QY |||||
QY 1441 GCATTTGGAGTGTGTGAGAAATGCTTAAAGATTTCTAATCAATGAAATTAAGGTTGACAAAC 1500
Db |||||
QY 1489 GCATTTGGAGTGTGTGAGAAATGCTTAAAGATTTCTAATCAATGAAATTAAGGTTGACAAAC 1548
QY |||||
QY 1501 AGCAGTATGGACGACCTTTCATAAGAAAGATGCTTGAATGTTTTCAGGCTCAAAGATGAACGT 1560
Db |||||
QY 1549 AGCAGTATGGACGACCTTTCATAAGAAAGATGCTTGAATGTTTTCAGGCTCAAAGATGAACGT 1608
QY |||||
QY 1561 GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
Db |||||
QY 1609 GACCTTGAAGATTTTCTGCTTGAATTTGAGGAAAGACCTGAAAGCCCTTCATTCAGTGCAG 1668
QY |||||
QY 1621 TGTTCACCTTCCCGAGGCCCTTCAAACTGTCGAACTGCTTGTATGGCTGCTGCTGATC 1680
Db |||||
QY 1669 TGTTCACCTTCCCGAGGCCCTTCAAACTGTCGAACTGCTTGTATGGCTGCTGCTGATC 1728
QY |||||
QY 1681 AGAATTTGAGTGTGGACCATAGCAGATTTCTGSCACTTACTTGTAAATGCTTTTGGTGACTTCA 1740
Db |||||
QY 1729 AGAATTTGAGTGTGGACCATAGCAGATTTCTGSCACTTACTTGTAAATGCTTTTGGTGACTTCA 1788
QY |||||
QY 1741 ACAGTTTTCAGATCCCTCTGTGTACATTTCCCTCCCAATTAACCTGTTAAATGGGGTCACTGCA 1800
Db |||||
QY 1789 ACAGTTTTCAGATCCCTCTGTGTACATTTCCCTCCCAATTAACCTGTTAAATGGGGTCACTGCA 1848
QY |||||
QY 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db |||||
QY 1849 GCAGTGAACATGCTCAGGGAGTCTCCAGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908
QY |||||
QY 1861 TTTGGCAGCTTTTGCAACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db |||||
QY 1909 TTTGGCAGCTTTTGCAACGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1968
QY |||||
QY 1921 GGTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 1980
Db |||||
QY 1969 GGTTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 2028
QY |||||
QY 1981 GAGCGTGGTGTCTGTGAAATATTTGCAAAATTTTGAACGAAAGCTTCCATTTTCTAGC 2040
Db |||||
QY 2029 GAGCGTGGTGTCTGTGAAATATTTGCAAAATTTTGAACGAAAGCTTCCATTTTCTAGC 2088
QY |||||
QY 2041 CTGAAAGTAATCAATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db |||||
QY 2089 CTGAAAGTAATCAATTTTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2148
QY |||||
QY 2101 CTGGGTGGCAGCAGATATGGGCTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db |||||
QY 2149 CTGGGTGGCAGCAGATATGGGCTTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2208
QY |||||
QY 2161 AGCACCATGGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db |||||
QY 2209 AGCACCATGGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2268
QY |||||
QY 2221 ACCAATTCCTACACCAAGCTCTACTGCAATTTGCAAGGGAGACCTTGGAATAATTTTGG 2280

QY 421 CAATCCCTCGTCTGATGCTAAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 480
DB 469 CAATCCCTCGTCTGATGCTAAACCAATCAGCTATGTGCCCCCAAGCTGTTTCAAGTGC 528
QY 481 CTGCATTCCTCAGGACCTGCTGGCTGGAATGCAATGCGTTAAACAGAAATCCCGTCCAG 540
DB 529 CTGCATTCCTCAGGACCTGCTGGCTGGAATGCAATGCGTTAAACAGAAATCCCGTCCAG 588
QY 541 GCTTTTGAAGTTTATCGCATTTGCAAGCATGACCTTGGCCCTGAAACAAAATACACCAAC 600
DB 589 GCTTTTGAAGTTTATCGCATTTGCAAGCATGACCTTGGCCCTGAAACAAAATACACCAAC 648
QY 601 ATACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTATCATCTCCATCAACAAT 660
DB 649 ATACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTATCATCTCCATCAACAAT 708
QY 661 AGAATCCACTCCTCGGGAAGAAATGCTTTGTATGGCTCCACAGCCTAGAGACTTTAGAT 720
DB 709 AGAATCCACTCCTCGGGAAGAAATGCTTTGTATGGCTCCACAGCCTAGAGACTTTAGAT 768
QY 721 TTAATTTACAATAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAA 780
DB 769 TTAATTTACAATAACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACTTAAA 828
QY 781 GAACTAGGATTTATAGCAACAAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 840
DB 829 GAACTAGGATTTATAGCAACAAATATCAGGTGATACCTGAGAAAGCAATTTGTAGGCAAC 888
QY 841 CCTCTCTTATTAACAATATCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
DB 889 CCTCTCTTATTAACAATATCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 948
QY 901 TTTCAACATTTTAACTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAAATGAA 960
DB 949 TTTCAACATTTTAACTGAACTAAGAACACTGACTCTGAATGGTGCCTCACAATAAATGAA 1008
QY 961 TTTCTGATTTTAACTGGAATCTGCAAACTGGAGAGCTGACTTTTAACTGGAGCAAGATC 1020
DB 1009 TTTCTGATTTTAACTGGAATCTGCAAACTGGAGAGCTGACTTTTAACTGGAGCAAGATC 1068
QY 1021 TCATCTCTCTCAACCGTCTGCAATCAGTTTACCTTAATCTCCAGTGTAGATCTGCT 1080
DB 1069 TCATCTCTCTCAACCGTCTGCAATCAGTTTACCTTAATCTCCAGTGTAGATCTGCT 1128
QY 1081 TACAACCTATTAGAAGATTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1140
DB 1129 TACAACCTATTAGAAGATTACCCAGTTTTCAGTCTGCCAAAAGCTTCAGAAAATTGAC 1188
QY 1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACATTTCCAGCAGTTGCTTAGCCTC 1200
DB 1189 CTAAGACATAATGAATCTACGAAATTAAGTTGACATTTCCAGCAGTTGCTTAGCCTC 1248
QY 1201 CGATCGCTGAATTTGGCTTGGAAACAAATGCTATTTATTCACCCCAATGCAATTTCCACT 1260
DB 1249 CGATCGCTGAATTTGGCTTGGAAACAAATGCTATTTATTCACCCCAATGCAATTTCCACT 1308
QY 1261 TTGCCATCCCTAATAAAGCTGACCTATCGTCAACCTCTGCTGCTTTTCCATAACT 1320
DB 1309 TTGCCATCCCTAATAAAGCTGACCTATGCTCAACCTCTGCTGCTTTTCCATAACT 1368
QY 1321 GGGTTACATGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1380
DB 1369 GGGTTACATGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA 1428
QY 1381 TCATCTGAAACTTTTCCAGAACTCAAGTTTATAGAAATGCTTATGCTTACAGGCTGT 1440
DB 1429 TCATCTGAAACTTTTCCAGAACTCAAGTTTATAGAAATGCTTATGCTTACAGGCTGT 1488
QY 1441 GCAATTTGGAGTGTGTGAGAAATGCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAA 1500
DB 1489 GCAATTTGGAGTGTGTGAGAAATGCTTATAGATTTCTTAATCAATGGAATAAAGGTGACAA 1548

QY 1501 AGCAGTATGGACGACCTTCATTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1560
DB 1549 AGCAGTATGGACGACCTTCATTAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACGT 1608
QY 1561 GACCTTTGAAGATTTTCTGCTTACATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1620
DB 1609 GACCTTTGAAGATTTTCTGCTTACATTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCAG 1668
QY 1621 TGTTCACTTCCCAAGGCCCTTCAAAACCTCTGTGAACACCTCTGTGAATGGCTGGCTGATC 1680
DB 1669 TGTTCACTTCCCAAGGCCCTTCAAAACCTCTGTGAACACCTCTGTGAATGGCTGGCTGATC 1728
QY 1681 AGAATTTGAGTGTGAGCAATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1740
DB 1729 AGAATTTGAGTGTGAGCAATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1788
QY 1741 ACAGATTTTTCAGATCCCTCTGTACATTTCCCCCACTTAAACTGTAAATGCGGTCAATCGCA 1800
DB 1789 ACAGATTTTTCAGATCCCTCTGTACATTTCCCCCACTTAAACTGTAAATGCGGTCAATCGCA 1848
QY 1801 GCAGTGAACATGCTCAACGGGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATGCGTTCACT 1860
DB 1849 GCAGTGAACATGCTCAACGGGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATGCGTTCACT 1908
QY 1861 TTTGGCAGCTTTGACAGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGCATGTCATT 1920
DB 1909 TTTGGCAGCTTTGACAGACATGCTGCTGGTGGGAGAAATGGGGTGGTGGCATGTCATT 1968
QY 1921 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTGCGAGCCCTG 1980
DB 1969 GGTTTTTTGTCCATTTTGTCTTCAAGATCATCTGTTTTCTGCTTACTCTGCGAGCCCTG 2028
QY 1981 GAGCGTGGGTTCCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
DB 2029 GAGCGTGGGTTCCTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2088
QY 2041 CTGAAGTAATCATTTTGTCTGTGCTGCTGCTGGCTTGACCATGGCCGAGTTCCCTCTG 2100
DB 2089 CTGAAGTAATCATTTTGTCTGTGCTGCTGCTGGCTTGACCATGGCCGAGTTCCCTCTG 2148
QY 2101 CTGGGTGGCAGCAAGTATGGGCTCCCTCTCTGCTGCTGCTGCTGCTTTTGGGAGCC 2160
DB 2149 CTGGGTGGCAGCAAGTATGGGCTCCCTCTCTGCTGCTGCTGCTTTTGGGAGCC 2208
QY 2161 AGCACCATTGGGCTACATGCTGCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2220
DB 2209 AGCACCATTGGGCTACATGCTGCTCATCTTGTCTCAATTCCTTTGCTTCTCATGATG 2268
QY 2221 ACCATGCTTACACCAAGCTCTACTGCAATTTGGACAAGGGAGAGCTTGAGAAATATTTGG 2280
DB 2269 ACCATGCTTACACCAAGCTCTACTGCAATTTGGACAAGGGAGAGCTTGAGAAATATTTGG 2328
QY 2281 GACTGCTTATGTTAAACACATTTGCCCTGTGCTCTTCCACCACTGCAATCTTAACTGC 2340
DB 2329 GACTGCTTATGTTAAACACATTTGCCCTGTGCTCTTCCACCACTGCAATCTTAACTGC 2388
QY 2341 CCTGTGGCTTTTCTGCTCTTCTCTTAAATAAACCCTTACATTTATCAGTCTGAGATA 2400
DB 2389 CCTGTGGCTTTTCTGCTCTTCTCTTAAATAAACCCTTACATTTATCAGTCTGAGATA 2448
QY 2401 APTAAGTTTATCTCTGCTGGTGTGCTCCACTTCTCTGCAATGCTCAATCCCTCTCTTAC 2460
DB 2449 APTAAGTTTATCTCTGCTGGTGTGCTCCACTTCTCTGCAATGCTCAATCCCTCTCTTAC 2508
QY 2461 ATCTGTTTCAATCTCTCACTTTTAAAGAGATCTGGTGGCTGAGAAAGCAACCTACGTC 2520
DB 2509 ATCTGTTTCAATCTCTCACTTTTAAAGAGATCTGGTGGCTGAGAAAGCAACCTACGTC 2568
QY 2521 TGGACAGATCAAAACACCCCAAGCTTGTGATGTTAACTCTGTGATGTTTCAAAAAACAG 2580
DB 2569 TGGACAGATCAAAACACCCCAAGCTTGTGATGTTAACTCTGTGATGTTTCAAAAAACAG 2628
QY 2581 TCCTGTGATCAACTCAAGCCTTGGTAAACCTTTTACCAGCTCCAGCATCACTTTATGACCTG 2640

Db 1669 TGTTCACCTTCCCGAGCCCTTCAAAACCCCTGGAACACACCTGCTTGATGGCTGGCTGATC 1728
QY
Db 1681 AGAATTGGAGTGGACCATAGCAGTTCTGGCACTTCTGTAATGCTTTGGTGACTTCA 1740
QY
Db 1729 AGAATTGGAGTGGACCATAGCAGTTCTGGCACTTCTGTAATGCTTTGGTGACTTCA 1788
QY
Db 1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTAAACCTGTAATTTGGGGTCATCGCA 1800
Db 1789 ACAGTTTTCAGATCCCTCTGTACATTTCCCACTAAACCTGTAATTTGGGGTCATCGCA 1848
QY
Db 1801 GCAGTGAACATGCTCAGGGAGTCTCAGTCCGCTGCTGGCTGGTGGATGGTTCACCT 1860
Db 1849 GCAGTGAACATGCTCAGGGAGTCTCAGTCCGCTGCTGGCTGGTGGATGGTTCACCT 1908
QY
Db 1861 TTGGGAGCTTTTGCAGCAGATGCTGCTGGTGGGAGAAATGGGGTGGTGGCCATGTCATT 1920
Db 1909 TTGGGAGCTTTTGCAGCAGATGCTGCTGGTGGGAGAAATGGGGTGGTGGCCATGTCATT 1968
QY
Db 1921 GGTGTTTGTGCAATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 1980
Db 1969 GGTGTTTGTGCAATTTTGTCTCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTG 2028
QY
Db 1981 GAGCGTGGTCTCTGTAATATCTGCAAAATTTGCAACGAAAGCTCAATTTCTAGC 2040
Db 2029 GAGCGTGGTCTCTGTAATATCTGCAAAATTTGCAACGAAAGCTCAATTTCTAGC 2088
QY
Db 2041 CTGAAAGTAAATCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db 2089 CTGAAAGTAAATCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148
QY
Db 2101 CTGGGTGGCAGCAAGTATGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db 2149 CTGGGTGGCAGCAAGTATGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2208
QY
Db 2161 AGCACCATGGGCTACATGGTGGTCTCATCTGCTCAATTCCTCTTGGCTTCTCATGATG 2220
Db 2209 AGCACCATGGGCTACATGGTGGTCTCATCTGCTCAATTCCTCTTGGCTTCTCATGATG 2268
QY
Db 2221 ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGATAATTTGG 2280
Db 2269 ACCATTGGCTTACACCAAGCTCTACTGCAATTTGGACAGGGAGACCTGGAGATAATTTGG 2328
QY
Db 2281 GACTGCTATAGTAAACACATGTCCTGCTCTTCAACCACTGCACTCTCAAACTGC 2340
Db 2329 GACTGCTATAGTAAACACATGTCCTGCTCTTCAACCACTGCACTCTCAAACTGC 2388
QY
Db 2341 CCTGTGGCTTCTGTGCTCTCTCTTTAAATAAACCCTTACATTTATCAGTCTGAAGTA 2400
Db 2389 CCTGTGGCTTCTGTGCTCTCTCTTTAAATAAACCCTTACATTTATCAGTCTGAAGTA 2448
QY
Db 2401 ATTAAGTTTATCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2449 ATTAAGTTTATCTCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2508
QY
Db 2461 ATCTGTTCAATCTCACTTTAAGGAGGATCTGGTGGAGCTTGAGAAAGCAAACTACGTC 2520
Db 2509 ATCTGTTCAATCTCACTTTAAGGAGGATCTGGTGGAGCTTGAGAAAGCAAACTACGTC 2568
QY
Db 2521 TGGAAGAATCAAAACACCAAGCTTGATGTCATTTAACTCTGATGATGTCGAAACAG 2580
Db 2569 TGGAAGAATCAAAACACCAAGCTTGATGTCATTTAACTCTGATGATGTCGAAACAG 2628
QY
Db 2581 TCTGTGACTCAACTCAAGCTTGGTAACTTTACCAGCTCCAGCATCACTTATGACCTG 2640
Db 2629 TCTGTGACTCAACTCAAGCTTGGTAACTTTACCAGCTCCAGCATCACTTATGACCTG 2688
QY
Db 2641 CCTCCAGTTCCGTGCCATCACCAGCTTATCCAGTGAAGTGGAGCTGCCATCTTTCTCT 2700
Db 2689 CCTCCAGTTCCGTGCCATCACCAGCTTATCCAGTGAAGTGGAGCTGCCATCTTTCTCT 2748
QY
Db 2701 GTGGCATTTGTCCCATGTCTCTAA 2724

Db 2749 GTGGCATTTGTCCCATGTCTCTAA 2772
RESULT 9
US-10-295-027-848
; Sequence 848, Application US/10295027
; Publication No. US2003023350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 848
; LENGTH: 3032
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-295-027-848
Query Match 100.0%; Score 2724; DB 17; Length 3032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2724; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGACACCTCCCGGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 201 ATGGACACCTCCCGGCTCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 260
QY 61 GGGGGCAGCTCTCCACAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 120
Db 261 GGGGGCAGCTCTCCACAGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 320
QY 121 GAGCCGACGGCAGGATGTTGCTCAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
Db 321 GAGCCGACGGCAGGATGTTGCTCAGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 380
QY 181 CTTTCCAACTCAGGCTCTTCACTCTACCTCTACCTCTACCTCTACCTCTACCTCTACCTCTACCT 240
Db 381 CTTTCCAACTCAGGCTCTTCACTCTACCTCTACCTCTACCTCTACCTCTACCTCTACCTCTACCT 440
QY 241 CTGCTCCCGAATCCCTCGCCAGTCTCGCTTCTCGGAGGAGTTACGTCTTGGGGAAAC 300

[illegible]

1081 TACACCATATTAGAGATTATCCCAAGTTTTCAGTCTGCCAAAGCTTCAGAAATTCAC 1140
Db TACAACTATTAGAGATTATCCCAAGTTTTCAGTCTGCCAAAGCTTCAGAAATTCAC 1340
1141 CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCACAGCAGTTGCTTAGCCTC 1200
Db CTAAGACATAATGAATCTACGAAATTAAGTTGACACTTTCACAGCAGTTGCTTAGCCTC 1400
1201 CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1260
Db CGATCGCTGAATTTGGCTTGGAAACAAATTTGCTATTATTCACCCCAATGCAATTTCCACT 1460
1261 TTGGCATCCCTAATAAGCTGACCTATGCTCCAACTCCCTGCTGCTTTTCCCTATAACT 1320
Db TTGGCATCCCTAATAAGCTGACCTATGCTCCAACTCCCTGCTGCTTTTCCCTATAACT 1520
1321 GGGTTACATGGTTTAACTCACCTTAAATTTAAACAGAAATCATGCTTACAGAGCTTGATA 1380
Db GGGTTACATGGTTTAACTCACCTTAAATTTAAACAGAAATCATGCTTACAGAGCTTGATA 1580
1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACAGTGCTGT 1440
Db TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTTACAGTGCTGT 1640
1441 GCATTTGGAGTGTGAGAAATGCCCTTAAAGATTTCTAATCAATGAATTAAGGTGACAAC 1500
Db GCATTTGGAGTGTGAGAAATGCCCTTAAAGATTTCTAATCAATGAATTAAGGTGACAAC 1700
1501 AGCAGTATGGAGCACTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAGT 1560
Db AGCAGTATGGAGCACTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAAGT 1760
1561 GACCTTGAAGATTTCTGCTTCACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCGAG 1620
Db GACCTTGAAGATTTCTGCTTCACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCGAG 1820
1621 TGTTTCACTTCCCGAGGCCCTTCAACCCCTGTGACACCTGCTGTGATGGCTGGCTGATC 1680
Db TGTTTCACTTCCCGAGGCCCTTCAACCCCTGTGACACCTGCTGTGATGGCTGGCTGATC 1880
1681 AGAATTGGAGTGTGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1740
Db AGAATTGGAGTGTGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA 1940
1741 ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAAATTTGGGGTCAATCGA 1800
Db ACAGTTTTCAGATCCCTCTGTACATTTCCCCCATTAACCTGTTAAATTTGGGGTCAATCGA 2000
1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTGGCTGCTGGCTGGTGGTGGTGGTGGTGGTGGT 1860
Db GCAGTGAACATGCTCAGGGAGTCTCCAGTGGCTGCTGGCTGGTGGTGGTGGTGGTGGTGGT 2060
1861 TTTGGCAGCTTTGCACGACATGGTGGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGT 1920
Db TTTGGCAGCTTTGCACGACATGGTGGCTGGTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGT 2120
1921 GGTGTTTTGTTCATTTTGTCTTCAAGATCATCTGTTTTCCTGCTTACTCTGTCAGCCCTG 1980
Db GGTGTTTTGTTCATTTTGTCTTCAAGATCATCTGTTTTCCTGCTTACTCTGTCAGCCCTG 2180
1981 GAGCGTGGGTTCTCTGTGAATTAATCTGCAAAATTTGAACCAAGCTCCATTTTCTAGC 2040
Db GAGCGTGGGTTCTCTGTGAATTAATCTGCAAAATTTGAACCAAGCTCCATTTTCTAGC 2240
2041 CTGAAAGTAAATATTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
Db CTGAAAGTAAATATTGCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2300
2101 CTGGGTGGCAGCAAGATATGGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2160
Db CTGGGTGGCAGCAAGATATGGGCGCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2360

2161 AGCACCATTGGCTACATGGTGGCTCTCATCTTGTCTCAATCCCTTTGCTTCTCATGATG 2220
Db AGCACCATTGGCTACATGGTGGCTCTCATCTTGTCTCAATCCCTTTGCTTCTCATGATG 2420
2221 ACCATTGGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGAGCTCGAGAAATTTGG 2280
Db ACCATTGGCTACACCAAGCTCTACTGCAATTTGGCAAGGGAGAGCTCGAGAAATTTGG 2480
2281 GACTGCTCTATGGTTAAACACATTTGCCCTGTTCTTCCACCACTGATCCTTAACTGC 2340
Db GACTGCTCTATGGTTAAACACATTTGCCCTGTTCTTCCACCACTGATCCTTAACTGC 2540
2341 CCTGGCTCTTCTTGTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2400
Db CCTGGCTCTTCTTGTCTCTCTCTTAAATAAACCTTACATTTATCAGTCTCGAAGTA 2600
2401 ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2460
Db ATTAAGTTTATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2660
2461 ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGAGCTGAGAAACCAACCTACGTC 2520
Db ATCTTGTTCATCTCTCACTTTAAGGAGATCTGGTGAGCTGAGAAACCAACCTACGTC 2720
2521 TGGCAAGATCAAAACACCCCAAGCTTGTATGTCAATTTAACTCTGTATGATGTGAAAAACAG 2580
Db TGGCAAGATCAAAACACCCCAAGCTTGTATGTCAATTTAACTCTGTATGATGTGAAAAACAG 2780
2581 TCCTGTGATCAACTCAAGCCTTGGTAACTTTTACCAGCTCAGCATCACTTATGACCTG 2640
Db TCCTGTGATCAACTCAAGCCTTGGTAACTTTTACCAGCTCAGCATCACTTATGACCTG 2840
2641 CTTCCAGCTTCCGTGCTCATCCAGCTTATCCAGTGTGAGAGCTGCTGCTTCTCTCT 2700
Db CTTCCAGCTTCCGTGCTCATCCAGCTTATCCAGTGTGAGAGCTGCTGCTTCTCTCTCT 2900
2701 GTGGCAATTTGCTCCATGCTCTAA 2724
Db GTGGCAATTTGCTCCATGCTCTAA 2924

RESULT 11

US-10-251-385-277
; Sequence 277, Application US/10251385
; Publication No. US20030105292A1
; GENERAL INFORMATION:
; APPLICANT: Behan, Dominic P.
; APPLICANT: Chalmers, Derek T.
; APPLICANT: Liaw, Chen W.
; TITLE OF INVENTION: No. US20030105292A1-Endogenous, Constitutively Activated Human G
; TITLE OF INVENTION: Protein-Coupled
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0040
; CURRENT APPLICATION NUMBER: US/10/251.385
; PRIOR FILING DATE: 2002-09-20
; PRIOR APPLICATION NUMBER: US/09/170,496
; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 294
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 277
; LENGTH: 2724
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-251-385-277

Query Match 99.9%; Score 2720.8; DB 15; Length 2724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2722; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 ATGGACACTCCCGGCTCGGTGTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 ATGGACACTCCCGGCTCGGTGTGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60

Qy		G G G G G C A G C T C C C A G G T T G G T G T G T G A G G G G C T G C C C C A C A C A C T G T C A T T G C	120
Db		G G G G G C A G C T C C C A G G T T G G T G T G T G T G A G G G G C T G C C C C A C A C A C T G T C A T T G C	120
Qy		G A G C C C G A G C G C A G S A N T T G C T C A G S G T G A C T G C T C G A C C T G S G G C T C T C G G A G C T G	180
Db		G A G C C C G A G C G C A G A T T G T C T C A G G T G A C T G C T C G G A C C T G G G G C T C T C G A G A C T G	180
Qy		C T T T C C A A C C T C A G C G T C T T C A C C T C C T A C C T A G A C C T C A G T A T G A A C A A C A T C A G T C A G	240
Db		C T T T C A A C C T C A G C G T C T T C A C C T C C T A C C T A G A C C T C A G T A T G A A C A A C A T C A G T C A G	240
Qy		C T G C T C C G G A A T C C C T G C C C A G T C T C G C T C C T G A G A G A T T A C G T C T T G C G G G A A A C	300
Db		C T G C T C C G G A A T C C C C T G C C C A G T C T C G C T C C T G A G A G A T T A C G T C T T G C G G G A A A C	300
Qy		G C T C T G A C A T A C A T T C C C A G G A G A C A T T C A C T G S C C T T T A C A G T C T T A A G T T C T T A T G	360
Db		G C T C T G A C A T A C A T T C C C A G G A G A C A T T C A C T G S C C T T T A C A G T C T T A A G T T C T T A T G	360
Qy		C T G C A G A A T A T C A G C T A A G A C A C G T A C C C A C A G A A G C T C T G C A G A A T T T G C G A A G C C T T	420
Db		C T G C A G A A T A T C A G C T A A G A C A C G T A C C C A C A G A A G C T C T G C A G A A T T T G C G A A G C C T T	420
Qy		C A A T C C C T G C G T G S A T G C T A A C A C A T C A G C T A T G T G C C C C A A G C T G T T C A G T G G C	480
Db		C A A T C C C T G C G T G S A T G C T A A C C A C A T C A G C T A T G T G C C C C A A G C T G T T C A G T G G C	480
Qy		C T G C A T T C C C T G A G C A C C T G T G C T G S A T G A C A A T G C G T T A A C A G A A A T C C C C G T C C A G	540
Db		C T G C A T T C C C T G A G C A C C T G T G C T G S A T G A C A A T G C G T T A A C A G A A A T C C C C G T C C A G	540
Qy		G C T T T T A G A A G T T T A T C G G C A T T C C A A G C C A T G A C C T T G G C C C T G A A C A A A A T A C A C C A C	600
Db		G C T T T T A G A A G T T T A T C G G C A T T C C A A G C C A T G A C C T T G G C C C T G A A C A A A A T A C A C C A C	600
Qy		A T A C C A G A T A T G C T T T T G G A A A C C T C T C A G C T T G T G A G T T C T A C A T C C A A T A A C A A T	660
Db		A T A C C A G A T A T G C T T T T G G A A A C C T C T C A G C T T G T G A G T T C T A C A T C C A A T A A C A A T	660
Qy		A G A A T C A C T C C C T G G G A A A G A A A T G C T T T G A T G S G C T C C A C A G C T A G A G A C T T T A G A T	720
Db		A G A A T C A C T C C C T G G G A A A G A A A T G C T T T G A T G S G C T C C A C A G C T A G A G A C T T T A G A T	720
Qy		T T A A A T T A C A A T A A C C T T G A T G A A T T C C C A C T G C A A T T A G G A C A C T C T C C A A C C T T A A A	780
Db		T T A A A T T A C A A T A A C C T T G A T G A A T T C C C A C T G C A A T T A G G A C A C T C T C C A A C C T T A A A	780
Qy		G A A C T A G A T T T C A T A G C A A C A A T A T C A G G F C G A T A C T G A G A A A G C A A T T T G T A G G C A A C	840
Db		G A A C T A G A T T T C A T A G C A A C A A T A T C A G G T C G A T A C T G A G A A A G C A A T T T G T A G G C A A C	840
Qy		C T T T C T C T A T T A C A A T A C A T T T C T A T A G C A A T C C C A T C C A A T T T G T G G G A G A T C T G C T	900
Db		C T T T C T C T A T T A C A A T A C A T T T C T A T A G C A A T C C C A T C C A A T T T G T G G G A G A T C T G C T	900
Qy		T T T C A A C A A T T T A C C T G A A C T A A G A A C A C T G A C T C T G A A T G G T G C T C A C A A A T A A C T G A A	960
Db		T T T C A A C A A T T T A C C T G A A C T A A G A A C A C T G A C T C T G A A T G G T G C T C A C A A A T A A C T G A A	960
Qy		T T T C C T G A T T P A A C T G G A A C T G C A A A C C T G S A G A G T C T G A C T T T P A A C T G G A G A C A G A T C	1020
Db		T T T C C T G A T T P A A C T G G A A C T G C A A A C C T G S A G A G T C T G A C T T T P A A C T G G A G A C A G A T C	1020
Qy		T C A T C T C T C T C C A A A C G G T C T G C A A T C A G T T A C T T A A T C C C A A G T C T A G A T C T G T C T	1080
Db		T C A T C T C T C T C C A A A C G G T C T G C A A T C A G T T A C T T A A T C C C A A G T C T A G A T C T G T C T	1080
Qy		T A C A A C C T A T T A G A A G A T T T A C C C A G T T T T T C A G T C T G C C A A A A G C T T C A G A A A T T G A C	1140
Db		T A C A A C C T A T T A G A A G A T T T A C C C A G T T T T T C A G T C T G C C A A A A G C T T C A G A A A T T G A C	1140
Qy		C T A A G A C A T A T G A A A T C T A C G A A A T T A A A G T T G A C A C A C T T T C C A G C A G T T G C T T A G C C T C	1200

1141	CTAAGACATAATGAAATCTACGAAATTAAGTTGACATTTCCAGACAGTTGCTTAGCCCTC	1201
1201	CGATCGCTGAATTTGGCTTGGAAACAAATTGCTATTATTACCCCCCAATGCAATTTCCACT	1261
1201	CGATCGCTGAATTTGGCTTGGAAACAAATTGCTATTATTACCCCCCAATGCAATTTCCACT	1261
1261	TTGCCATCCCTAATAAAGCTGGACCTTATCGTCCAACTCGTGGTGGTCTTTCCTATAACT	1321
1261	TTGCCATCCCTAATAAAGCTGGACCTTATCGTCCAACTCGTGGTGGTCTTTCCTATAACT	1321
1321	GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA	1381
1321	GGGTTACATGGTTTAACTCACTTAAATTAACAGGAAATCATGCTTACAGAGCTTGATA	1381
1381	TCATCTGMAAATCTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGCCTG	1441
1381	TCATCTGMAAATCTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGCCTG	1441
1441	GCATTTGGAGTGTGAGAAATGCTTATAAGATTTCTAATCAATTGGAAATAAGGTGACAAC	1501
1441	GCATTTGGAGTGTGAGAAATGCTTATAAGATTTCTAATCAATTGGAAATAAGGTGACAAC	1501
1501	AGCAGTATGACACACTTTCATATAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT	1561
1501	AGCAGTATGACACACTTTCATATAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAACGT	1561
1561	GACCTTGAGATTTCTGCTTGACTTTTGAGGAAGACTGAAAGCCCTTCATTTCAGTGCAG	1621
1561	GACCTTGAGATTTCTGCTTGACTTTTGAGGAAGACTGAAAGCCCTTCATTTCAGTGCAG	1621
1621	TGTTCACTTTCCCAGGCCCCCTTCAAAACCTCTGGAACACTGCTTGATGCTGGCTGATC	1681
1621	TGTTCACTTTCCCAGGCCCCCTTCAAAACCTCTGGAACACTGCTTGATGCTGGCTGATC	1681
1681	AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1741
1681	AGAATTTGGAGTGTGGACCATAGCAGTTCTGGCACTTACTTGTAAATGCTTTGGTGACTTCA	1741
1741	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAATGTTAAATGGGGTCATCGCA	1801
1741	ACAGTTTTTCAGATCCCTCTGTACATTTCCCCCAATTAATGTTAAATGGGGTCATCGCA	1801
1801	GCAGTGAACAATGCTCACGGAGTCTCCAGTGCCGCTGCTGGCTGGTGGATGCGTTCACT	1861
1801	GCAGTGAACAATGCTCACGGAGTCTCCAGTGCCGCTGCTGGCTGGTGGATGCGTTCACT	1861
1861	TTTGGCAGCTTTGCACACATATGCTGCTCAGAAATCATCTGTTTTCTCTGCTTACTTGGCCATGCAATT	1921
1861	TTTGGCAGCTTTGCACACATATGCTGCTCAGAAATCATCTGTTTTCTCTGCTTACTTGGCCATGCAATT	1921
1921	GGTTTTTTGTCCAATTTTGCTTCAGAAATCATCTGTTTTCTCTGCTTACTTGGCAGCCCTG	1981
1921	GGTTTTTTGTCCAATTTTGCTTCAGAAATCATCTGTTTTCTCTGCTTACTTGGCAGCCCTG	1981
1981	GAGCGTGGGTTCTCTGTGAAATATTTCTGAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2041
1981	GAGCGTGGGTTCTCTGTGAAATATTTCTGAAAATTTGAAACGAAAGCTCCATTTTCTAGC	2041
2041	CTGAAAGTAAATCATTTTGTCTGTGCCCCTGCTGGCCCTTGACCATGCGCGAGTTCCCTG	2101
2041	CTGAAAGTAAATCATTTTGTCTGTGCCCCTGCTGGCCCTTGACCATGCGCGAGTTCCCTG	2101
2101	CTGGGTGGCAGCAAGTATGCGCCTCCCTCTCTGCTGCTGCTTTTGGCTTTGGGAGCCCC	2161
2101	CTGGGTGGCAGCAAGTATGCGCCTCCCTCTCTGCTGCTGCTTTTGGCTTTGGGAGCCCC	2161
2161	AGCACCATGGGTCATGCTGCTCATCTTGTCCAAATTCCTTTGCTTCTCTCATGATG	2221
2161	AGCACCATGGGTCATGCTGCTCATCTTGTCCAAATTCCTTTGCTTCTCTCATGATG	2221
2221	ACCATTGCCTPACCAAGCTCTACTGCAATTTGGACAAGGAGACTGGAGAAATATTTGG	2281

Qy	1261	TTGGCATCCCTTAATAAAGCTGGGACCTATAGCTGCACAACTCCTCTGTCGGTCTTTTCTTAATAACT	1320
Db	1261	TTGGCATCCCTTAATAAAGCTGGGACCTATAGCTGCACAACTCCTCTGTCGGTCTTTTCTTAATAACT	1320
Qy	1321	GGGTTACATGCTTTAACTCACTTAAATAATTAACAGGAAATCATGCTCTTACAGAGCTTGATA	1380
Db	1321	GGGTTACATGCTTTAACTCACTTAAATAATTAACAGGAAATCATGCTCTTACAGAGCTTGATA	1380
Qy	1381	TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGCCTGT	1440
Db	1381	TCATCTGAAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGCCTGT	1440
Qy	1441	GCATTTGGAGTGTGTGAGATGCTTAAAGATTTCTAATTCOAATGGAATAAAGGTGACAAC	1500
Db	1441	GCATTTGGAGTGTGTGAGATGCTTAAAGATTTCTAATTCOAATGGAATAAAGGTGACAAC	1500
Qy	1501	AGCAGTATGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560
Db	1501	AGCAGTATGACGACCTTCATAAGAAAGATGCTGGAATGTTTCAGGCTCAAGATGAACGT	1560
Qy	1561	GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCGAG	1620
Db	1561	GACCTTGAAGATTTCTGCTTGACTTTGAGGAAGACCTGAAAGCCCTTCATTCAGTGCGAG	1620
Qy	1621	TGTTCACTTCCCGCAGGCCCTTCAAAACCTGTGAAACCTGCTTGAATGCTGGCTGATC	1680
Db	1621	TGTTCACTTCCCGCAGGCCCTTCAAAACCTGTGAAACCTGCTTGAATGCTGGCTGATC	1680
Qy	1681	AGAAATGGAGTGTGGACCATAGCAGTCTTGCGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Db	1681	AGAAATGGAGTGTGGACCATAGCAGTCTTGCGCACTTACTTGTAAATGCTTTGGTGACTTCA	1740
Qy	1741	ACAGTTTTCAGATCCCTCTGTGACATTTCCCCCAATTAACCTGTTAAATGGGGTCATCGCA	1800
Db	1741	ACAGTTTTCAGATCCCTCTGTGACATTTCCCCCAATTAACCTGTTAAATGGGGTCATCGCA	1800
Qy	1801	GCAGTGAACATGCTCA CGGGAGTCTC CAGTGCCTGTGCTGGTGGTGAATGCGTTCACT	1860
Db	1801	GCAGTGAACATGCTCA CGGGAGTCTC CAGTGCCTGTGCTGGTGGTGAATGCGTTCACT	1860
Qy	1861	TTTTGGCAGCTTTGCACGACATGGTGCCTGGTGGGAGAAATGGGTGGTGGCCATGTCATT	1920
Db	1861	TTTTGGCAGCTTTGCACGACATGGTGCCTGGTGGGAGAAATGGGTGGTGGCCATGTCATT	1920
Qy	1921	GGTTTTTGTGCCATTTTTTGTCTCAGAAATCATCTGTTTTCTCTGCTTACTCTGGCAGCCCTG	1980
Db	1921	GGTTTTTGTGCCATTTTTTGTCTCAGAAATCATCTGTTTTCTCTGCTTACTCTGGCAGCCCTG	1980
Qy	1981	GAGCGTGGGTTCTCTGTAANAATTTCTGCAAAATTTGAAAACGAAAGCTCAATTTTCTPAGC	2040
Db	1981	GAGCGTGGGTTCTCTGTAANAATTTCTGCAAAATTTGAAAACGAAAGCTCAATTTTCTPAGC	2040
Qy	2041	CTGAAAGTAAATCATTTTGTCTCTGTCCTCTGTCCTGCTTGAACCATGCGCGCAGTTCGCCCTG	2100
Db	2041	CTGAAAGTAAATCATTTTGTCTCTGTCCTCTGTCCTGCTTGAACCATGCGCGCAGTTCGCCCTG	2100
Qy	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGTCCTGCTTTCCTTTTGGGAGCCC	2160
Db	2101	CTGGGTGGCAGCAAGTATGGCGCTCCCTCTCTGTCCTGCTTTCCTTTTGGGAGCCC	2160
Qy	2161	AGCACCATGGGCTACATGGTGCCTCTCATCTTTGCTCAATTTCCCTTTGCTTCTCATGATG	2220
Db	2161	AGCACCATGGGCTACATGGTGCCTCTCATCTTTGCTCAATTTCCCTTTGCTTCTCATGATG	2220
Qy	2221	ACCAATTGGCTACACCAAGCTCTACTGCAATTTTGAACAAGGAGACCTCGGAGAAATATTGG	2280
Db	2221	ACCAATTGGCTACACCAAGCTCTACTGCAATTTTGAACAAGGAGACCTCGGAGAAATATTGG	2280
Qy	2281	GACTGCTCTATGGTAAACAACATTTGCCCTCTTCTTCTTCCAACTGCATCTCAAACCTGC	2340
Db	2281	GACTGCTCTATGGTAAACAACATTTGCCCTCTTCTTCTTCCAACTGCATCTCAAACCTGC	2340
Qy	2341	CCTGTGCTCTTCTGTGCTCTCTCTCTTTTAATAAACCTTACATTTATTCAGTCTCTCAAGTA	2400

RESULT. T 13

REF ID: A63011 13
US-10-295-027-483

US-10-233-027-463
: Sequence 483. Application US/10295027

; sequence 483, Application US/10
: Publication No. US20030232350A1

: PUBLICATION NO. US20

APPLICANT: Afar Daniel

;; APPLICANT: ALAR, DANIEL
: APPLICANT: AZIZ, NATASHA

APPLICANT: Aziz, Natasha
APPLICANT: Cinsborg Wendy M

APPLICANT: Ginsberg, Wendy
APPLICANT: Gish, Kurt C

APPLICANT: Gish, Kurt C.
APPLICANT: Clinton Pishard

APPLICANT: Glynn, Richard

APPLICANT: Hevezi, Peter A.

APPLICANT: Mack, David H.

; APPLICANT: Murray, Richard

APPLICANT: Watson, Susan R.

APPLICANT: Eos Biotechnology, Inc.

; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions a

; TITLE OF INVENTION: Methods of

; FILE REFERENCE: 018501-012500US

; CURRENT APPLICATION NUMBER: US/1

; CURRENT FILING DATE: 2002-11-13

; PRIOR APPLICATION NUMBER: US 09/663,733

PRIOR FILING DATE: 2000-09-15

; PRIOR APPLICATION NUMBER: US 60/350,666

PRIOR FILING DATE: 2001-11-13

: PRIOR APPLICATION NUMBER: US 60/335,394

: PRIOR FILING DATE: 2001-11-15
 , ERIK AFFIDAVIT NUMBER: 03 00/333,334

; PRIORITY FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/332,464

; PRIOR APPLICATION NUMBER: US 80/332,484
 : PRIOR FILING DATE: 2001-11-21

; PRIOR FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: US 60/334 393

;; PRIOR APPLICATION NUMBER: US 60/334,393
: PRIOR FILING DATE: 2001-11-29

; PRIOR FILING DATE: 2001-11-29
 ; PRIOR APPLICATION NUMBER: US 60/340 376

;; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14

; PRIOR FILING DATE: 2001-12-14
 ; PRIOR APPLICATION NUMBER: US 60/347 311

; PRIOR APPLICATION NUMBER: US 60/347,211

; PRIOR FILING DATE: 2002-01-08
 ; PRIOR APPLICATION NUMBER: US 60/345 340

; PRIOR APPLICATION NUMBER: US 60/347,349

; PRIOR FILING DATE: 2002-01-10

; PRIOR APPLICATION NUMBER: US 60/355,250

; PRIOR FILING DATE: 2002-02-08

; PRIOR APPLICATION NUMBER: US 60/356,714

; PRIOR FILING DATE: 2002-02

; Remaining Prior Application

; NUMBER OF SEQ

; SOFTWARE: Pat

Db 661 AGAATCCATCCCTCGGGAAGAAATGCTTTTGATGGCTCCACAGCTAGAGACTTTAGAT 720
Qy 721 TTAATTTACAATAAAGCTTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
Db 721 TTAATTTACAATAAAGCTTTGATGAATTTCCCACTGCAATTTAGGACACTCTCCAACTTAAA 780
Qy 781 GAACCTAGGATTTTCATAGCAACAATATCAGGTGATACCTGAGAAAGCATTTGTAGGCAAC 840
Db 781 GAAC----- 784
Qy 841 CTTCTCTTTATTAACAATATTTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 785 -----TACATTTCTATGACAAATCCCATCCCAATTTGTTGGGAGATCTGCT 828
Qy 901 TTTCAACATTTACCTGAACCTAGAACACCTGACTCTGAATGGTGCCTTCACAAATTAATGAA 960
Db 829 TTTCAACATTTACCTGAACCTAGAACACCTGACTCTGAATGGTGCCTTCACAAATTAATGAA 888
Qy 961 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 1020
Db 889 TTTCTCTGATTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTAACTGGAGCACAGATC 948
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAAATCTCCAAAGTGTAGATCTGCT 1080
Db 949 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTTACCTAAATCTCCAAAGTGTAGATCTGCT 1008
Qy 1081 TACAACCTATTAGAGATTTACCAAGTTTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC 1140
Db 1009 TACAACCTATTAGAGATTTACCAAGTTTTTTCAGTCTGCGCAAAAGCTTCAGAAATTTGAC 1068
Qy 1141 CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTCCAGCAGTTGCTTAGCCCTC 1200
Db 1069 CTAAGACATAATGAAATCTACGAAATTTAAAGTTTGACACTTTCCAGCAGTTGCTTAGCCCTC 1128
Qy 1201 CGATCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTTACCCCAATGCAATTTTCCACT 1260
Db 1129 CGATCGCTGAATTTGGCTTGGAAACAAAATGCTATTATTTACCCCAATGCAATTTTCCACT 1188
Qy 1261 TTGCCATCGCTTAATTAAGCTGACCTATGCTCAACCTCTGCTCTTTTCTCTATAACT 1320
Db 1189 TTGCCATCGCTTAATTAAGCTGACCTATGCTCAACCTCTGCTCTTTTCTCTATAACT 1248
Qy 1321 GGGTTTACATGGTTTAACTACACTTAAATTTAAACAGAAATCATGCTTACAGAGCTTGATA 1380
Db 1249 GGGTTTACATGGTTTAACTACACTTAAATTTAAACAGAAATCATGCTTACAGAGCTTGATA 1308
Qy 1381 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGTGTGT 1440
Db 1309 TCATCTGAAACCTTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACAGTGTGT 1368
Qy 1441 GCATTTGGAGTGTGAGNAATGCCCTATAAGATTTCTTAATCAATGGAATAAAGGTGACAAAC 1500
Db 1369 GCATTTGGAGTGTGAGNAATGCCCTATAAGATTTCTTAATCAATGGAATAAAGGTGACAAAC 1428
Qy 1501 AGCAGTATGAGAGCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAACT 1560
Db 1429 AGCAGTATGAGAGCTTCATAGAAAGATGCTGGAATGTTTCAGGCTCAAGATCAACT 1488
Qy 1561 GACCTTGAAGATTTCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1620
Db 1489 GACCTTGAAGATTTCTGCTGACTTTGAGGAAGACCTGAAAGCCCTTCAATTCAGTGCAG 1548
Qy 1621 TGTTCACCTTCCAGAGGCCCTTCAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC 1680
Db 1549 TGTTCACCTTCCAGAGGCCCTTCAACCCCTGTGAACACCTGCTTGATGGCTGGCTGATC 1608
Qy 1681 AGAATTTGGAGTGTGGACCAATAGACTTCTGGCACTTACTTTGTAATGCTTTGTGACTTCA 1740
Db 1609 AGAATTTGGAGTGTGGACCAATAGACTTCTGGCACTTACTTTGTAATGCTTTGTGACTTCA 1668
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTCCCAATTTAAACTGTTAAATTTGGGGTCAATCGCA 1800
Db 1669 ACAGTTTTTCAGATCCCTCTGTACATTTCCCAATTTAAACTGTTAAATTTGGGGTCAATCGCA 1728

Qy 1801 GCAGTGAACATGCTTCAACGGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATCGGTTCACT 1860
Db 1729 GCAGTGAACATGCTTCAACGGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATCGGTTCACT 1788
Qy 1861 TTTTGCAGCTTTTGACGACATGCTGCTGGTGGGGAATAGGGTGTGGTTCATGTCATTT 1920
Db 1789 TTTTGCAGCTTTTGACGACATGCTGCTGGTGGGGAATAGGGTGTGGTTCATGTCATTT 1848
Qy 1921 GGTTTTTTGTGCAATTTTGTGCTTCAAGATCATCTGTTTTCTGCTTACTCTTGCAGCCCTG 1980
Db 1849 GGTTTTTTGTGCAATTTTGTGCTTCAAGATCATCTGTTTTCTGCTTACTCTTGCAGCCCTG 1908
Qy 1981 GAGCGTGGTGTCTCTGTGCAAAATATTTCTGCAAAATTTTGAAGGAAAGCTCCATTTTCTAGC 2040
Db 1909 GAGCGTGGTGTCTCTGTGCAAAATATTTCTGCAAAATTTTGAAGGAAAGCTCCATTTTCTAGC 1968
Qy 2041 CTGAAGTAATCAATTTTGTCTCTGTGCTGCTGCTGCTTGGACCATGGCCGAGTTCCCTCTG 2100
Db 1969 CTGAAGTAATCAATTTTGTCTCTGTGCTGCTGCTGCTTGGACCATGGCCGAGTTCCCTCTG 2028
Qy 2101 CTGGTGGGAGAGTATGGGCTCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGCCCC 2160
Db 2029 CTGGTGGGAGAGTATGGGCTCTCCCTCTCTGCTGCTGCTTGGCTTTTGGGGAGCCCC 2088
Qy 2161 AGCACATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTCTTGTCTTCTCATGATG 2220
Db 2089 AGCACATGGGCTACATGCTGCTCTCATCTTGTCTCAATTCCTCTTGTCTTCTCATGATG 2148
Qy 2221 ACCATTGCTTACACCAAGCTCTACTGCAATTTTGGCAAGGGAGACCTGGAGAAATTTTGG 2280
Db 2149 ACCATTGCTTACACCAAGCTCTACTGCAATTTTGGCAAGGGAGACCTGGAGAAATTTTGG 2208
Qy 2281 GACTGCTCTATGGTAAACACATTTGCCCTTGTGCTTCTTCAACCACTGATCCTTAAACTGTC 2340
Db 2209 GACTGCTCTATGGTAAACACATTTGCCCTTGTGCTTCTTCAACCACTGATCCTTAAACTGTC 2268
Qy 2341 CCTGTGGCTTTTCTGCTGCTTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTCTCAAGTA 2400
Db 2269 CCTGTGGCTTTTCTGCTGCTTCTCTCTTTTAAATAAACCTTACATTTATCAGTCTCTCAAGTA 2328
Qy 2401 ATTAAGTTTATCTCTTCTGGTGGTATGCCACTTCTGCTGATGCTCAATTCCTCTCTAC 2460
Db 2329 ATTAAGTTTATCTCTTCTGGTGGTATGCCACTTCTGCTGATGCTCAATTCCTCTCTAC 2388
Qy 2461 ATCTTGTTCATCTCTCACTTTTAAAGAGGATCTGGTGGGCTGAGAAAGCAAACTTACGTC 2520
Db 2389 ATCTTGTTCATCTCTCACTTTTAAAGAGGATCTGGTGGGCTGAGAAAGCAAACTTACGTC 2448
Qy 2521 TGGCAAGATCAAAACACCCCAAGCTTGTGATGTCAATTTAACTCTGTGATGATGTCGAAAAACAG 2580
Db 2449 TGGCAAGATCAAAACACCCCAAGCTTGTGATGTCAATTTAACTCTGTGATGATGTCGAAAAACAG 2508
Qy 2581 TCCTGTGACTCAACTCAAGCTTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTG 2640
Db 2509 TCCTGTGACTCAACTCAAGCTTTGGTAACTTTTACCAGCTCCAGCATCACTTATGACCTG 2568
Qy 2641 CCTCCAGTTCGGTGCATCACCAGCTTATCCAGTGTGAGAGCTGCCATCTTTCCTCT 2700
Db 2569 CCTCCAGTTCGGTGCATCACCAGCTTATCCAGTGTGAGAGCTGCCATCTTTCCTCT 2628
Qy 2701 GTGGCATTTGTCCCATGTCT 2720
Db 2629 GTGGCATTTGTCCCATGTCT 2648

RESULT 15

US-10-173-999-27
; Sequence 27, Application US/10173999
; Publication No. US20040005563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.


```
Qy 1801 GCAGTGAACATGCTCAGGGAGTCTCCAGTCCGTGCTGGTGTGGATGGTTCACCT 1860
Db |||||
Qy 1729 GCAGTGAACATGCTCAGGGAGTCTCCAGTCCGTGCTGGTGTGGATGGTTCACCT 1788
Db |||||
Qy 1861 TTTGGCAGCTTTTGACCGACATGGTGCCTGCTGGGAGAAATGGGGTTGGTGCATGTCAAT 1920
Db |||||
Qy 1789 TTTGGCAGCTTTTGACCGACATGGTGCCTGCTGGGAGAAATGGGGTTGGTGCATGTCAAT 1848
Db |||||
Qy 1921 GGTTTTGTGTCATTTTGTCTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTG 1980
Db |||||
Qy 1849 GGTTTTGTGTCATTTTGTCTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTG 1908
Db |||||
Qy 1981 GAGCGTGGGTCTCTGTGAATATTTCTGAAATATTTGAAACGAAAGCTCCCAATTTCTAGC 2040
Db |||||
Qy 1909 GAGCGTGGGTCTCTGTGAATATTTCTGAAATATTTGAAACGAAAGCTCCCAATTTCTAGC 1968
Db |||||
Qy 2041 CTGAAAGTAATCATTTTGTCTCTGTGCCCTGTGCTGCTTGACCATGGCCGAGTTCCCTCG 2100
Db |||||
Qy 1969 CTGAAAGTAATCATTTTGTCTCTGTGCCCTGTGCTGCTTGACCATGGCCGAGTTCCCTCG 2028
Db |||||
Qy 2101 CTGGGTGGCAGAGATATGGGCGCTCCCTCTCTGCTGCTGCTTTGGGAGGCC 2160
Db |||||
Qy 2029 CTGGGTGGCAGAGATATGGGCGCTCCCTCTCTGCTGCTGCTTTGGGAGGCC 2088
Db |||||
Qy 2161 AGCACCATGGGTCTCATGCTCGCTCATCTTGTCTCAATTCCTTGTCTCTCATGATG 2220
Db |||||
Qy 2089 AGCACCATGGGTCTCATGCTCGCTCATCTTGTCTCAATTCCTTGTCTCTCATGATG 2148
Db |||||
Qy 2221 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAAGGAGAGCTGGAGAAATTTGG 2280
Db |||||
Qy 2149 ACCATTGCTACACCAAGCTCTACTGCAATTTGGACAAAGGAGAGCTGGAGAAATTTGG 2208
Db |||||
Qy 2281 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTACCAAACTGATCCTAAACTGC 2340
Db |||||
Qy 2209 GACTGCTCTATGGTAAACACATTTGCCCTGTGCTCTTACCAAACTGATCCTAAACTGC 2268
Db |||||
Qy 2341 CCTGTGGCTTTCTGTCTCTCTCTTTAATAAACCTTTACATTTATCAGTCCCTGAAGTA 2400
Db |||||
Qy 2269 CCTGTGGCTTTCTGTCTCTCTCTTTAATAAACCTTTACATTTATCAGTCCCTGAAGTA 2328
Db |||||
Qy 2401 ATTAAGTTTATCCTTTCTGGTGGTAGTCCCACTTCTGCAATGTCTCAATCCCCTTCTCTAC 2460
Db |||||
Qy 2329 ATTAAGTTTATCCTTTCTGGTGGTAGTCCCACTTCTGCAATGTCTCAATCCCCTTCTCTAC 2388
Db |||||
Qy 2461 ATCTTGTTCATCTCCTTTAAGAGGATCTGGTAGCCTGAGAAAGCAACCTACGTC 2520
Db |||||
Qy 2389 ATCTTGTTCATCTCCTTTAAGAGGATCTGGTAGCCTGAGAAAGCAACCTACGTC 2448
Db |||||
Qy 2521 TGGACAAGATCAAAACACCCCAAGCTTGATGTCATTTAACTCTGATGATGTCGAAAAACAG 2580
Db |||||
Qy 2449 TGGACAAGATCAAAACACCCCAAGCTTGATGTCATTTAACTCTGATGATGTCGAAAAACAG 2508
Db |||||
Qy 2581 TCCTGTGACTCAACTCAAGCCTTGATTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2640
Db |||||
Qy 2509 TCCTGTGACTCAACTCAAGCCTTGATTAACCTTTTACAGCTCCAGCATCACTTATGACCTG 2568
Db |||||
Qy 2641 CCTCCAGTTCGGTGCATCAACAGCTTATCCAGTACTGAGAGCTGCCATCTTTCCTCT 2700
Db |||||
Qy 2569 CCTCCAGTTCGGTGCATCAACAGCTTATCCAGTACTGAGAGCTGCCATCTTTCCTCT 2628
Db |||||
Qy 2701 GTGGCATTTGTCCCATGTCT 2720
Db |||||
Qy 2629 GTGGCATTTGTCCCATGTCT 2648
Db |||||
```

Search completed: July 12, 2005, 21:05:45
Job time : 1831.02 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2652	97.4	2724	9	AY411732	Homo sapi
2	2004.6	73.6	2724	9	AY411734	Mus muscu
3	1958	71.9	2593	9	AY411733	Pan trogl
4	1037	38.1	1104	1	AL530798	AL530798
5	979.2	35.9	1060	1	AL541959	AL541959
6	845.8	31.0	4796	3	AK047873	Mus muscu
7	805.8	29.6	3583	3	AK085901	Mus muscu
8	745.2	27.4	940	7	CK449409	892576 MA
9	738.6	27.1	800	4	BM796708	RST15934
10	669.4	24.6	671	4	BM795259	K-EST0077
11	661.4	24.3	3509	3	BC038795	Homo sapi
12	604.4	22.2	763	7	CF742674	UI-M-HB0
13	585.4	21.5	873	5	BX403219	BX403219
14	576.4	21.2	3035	3	AK044357	Mus muscu
15	573.4	21.0	590	6	CD617354	55041415H
16	524	19.2	2822	3	AK052873	Mus muscu
17	521.2	19.1	776	5	BU708439	UI-M-FC0
18	500.8	18.4	639	3	AI697103	tg14f08.x
19	486.4	17.9	695	6	BY731701	BY731701
20	484.2	17.8	634	7	CF177761	806371 MA
21	475	17.4	475	5	BX107244	EX107244
22	472	17.3	472	1	AA460529	zx6809.r
23	471	17.3	494	1	AA424098	zv80c01.r
24	467.4	17.2	556	5	BQ323949	IL5-CI014

Db 1 ATGGACACCTCCCGGCTCGGTGTGCTCTGTCTCTTGTCCCTGTGCTGCTGCAGCTGGCGACC 60
Qy 61 GGGGCGAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCCAACACATGTCATTGC 120
Db 61 GGGGCGAGCTCTCCAGGCTCTGGTGTGTTGCTGAGGGGCTGCCCAACACATGTCATTGC 120
Qy 121 GAGCCCGAGCGGAGGATGTTGCTCAGGCTGGAAGTCTCCGACCTGGGCTCTCGGAGCTG 180
Db 121 GAGCCCGAGCGGAGGATGTTGCTCAGGCTGGAAGTCTCCGACCTGGGCTCTCGGAGCTG 180
Qy 181 CTTTCCAACTCAGCGCTCTTCACTCTCCCTACCTAGACCTCAGTATGAAACAAATCAGTCAG 240
Db 181 CTTTCCAACTCAGCGCTCTTCACTCTCCCTACCTAGACCTCAGTATGAAACAAATCAGTCAG 240
Qy 241 CTGCTCCCGAATCCCTCGCCAGTCTCCGCTTCTGAGGAGTACGTCCTTGGCGGAAAC 300
Db 241 CTGCTCCCGAATCCCTCGCCAGTCTCCGCTTCTGAGGAGTACGTCCTTGGCGGAAAC 300
Qy 301 GCTCTGACATACATTTCCCAAGGAGCATTCACCTGGCCCTTACAGTCTTTAAAGTCTTATG 360
Db 301 GCTCTGACATACATTTCCCAAGGAGCATTCACCTGGCCCTTACAGTCTTTAAAGTCTTNNN 360
Qy 361 CTGCAGAAATAACAGCTAAGACACGTAACCCACAGAAAGCTCTGCAGAAATTTGCGAAGCCTT 420
Db 361 NNN 420
Qy 421 CAATCCCTCGCTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Db 421 NNNNNNNNGCGCTCTGGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGC 480
Qy 481 CTGCATTCCTTGAGGACACTGTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Db 481 CTGCATTCCTTGAGGACACTGTGGCTGGATGACAAATGCGTTAAACAGAAATCCCGTCCAG 540
Qy 541 GCTTTTGAAGTTTATCGGCATTCGACATTCGAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
Db 541 GCTTTTGAAGTTTATCGGCATTCGAGCCATGACCTTTGGCCCTGAAACAAATACACCAC 600
Qy 601 ATACAGACTATGCTTTTGGAACTCTCCAGCTTGGTGTAGTCTTACATCTCCATACAAAT 660
Db 601 ATACAGACTATGCTTTTGGAACTCTCCAGCTTGGTGTAGTCTTACATCTCCATACAAAT 660
Qy 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Db 661 AGAATCCACTCCCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCTTAGAGACTTTAGAT 720
Qy 721 TTAATATCAATAACCTTTGATGAAATTCGCCACTGCAATTTAGGACACTCTCCAAACCTTAA 780
Db 721 TTAATATCAATAACCTTTGATGAAATTCGCCACTGCAATTTAGGACACTCTCCAAACCTTAA 780
Qy 781 GAACTAGGATTTATAGCAACAATATCAGTTCGATACCTGGAAGCAATTTGTAGGCAAC 840
Db 781 GAACTAGGATTTATAGCAACAATATCAGTTCGATACCTGGAAGCAATTTGTAGGCAAC 840
Qy 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Db 841 CCTTCTCTTATTAACAATACATTTCTATGACAAATCCATCCCAATTTGTTGGGAGATCTGCT 900
Qy 901 TTTCAACAATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTCTCAAAATAACTGAA 960
Db 901 TTTCAACAATTTACCTGAACTAAGAACACTGACTCTGAAATGGTGGCTCTCAAAATAACTGAA 960
Qy 961 TTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Db 961 TTTCTGATTTTAACTGGAACTGCAAACTGGAGAGTCTGACTTTTAACTGGAGCAAGATC 1020
Qy 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTCATTAATCTCCAAAGTGTAGATCTGTCT 1080
Db 1021 TCATCTCTTCTCAAAACCGTCTGCAATCAGTTCATTAATCTCCAAAGTGTAGATCTGTCT 1080
Qy 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAGCTTTCAGAAATTCAG 1140
Db 1081 TACAACCTATTAGAAGATTTACCCAGTTTTTCAGTCTGCCAAAGCTTTCAGAAATTCAG 1140

Qy 1141 CTAAGACATATGAAATCTACGAAATTAAGTTTGACACCTTTCCAGCAGTGTGCTTAGCCTC 1200
Db 1141 CTAAGACATATGAAATCTACGAAATTAAGTTTGACACCTTTCCAGCAGTGTGCTTAGCCTC 1200
Qy 1201 CGATCGCTGAATTTGGCTTTGGAAACAAATTTCTATTATTACCCCAATGCAATTTTCCACT 1260
Db 1201 CGATCGCTGAATTTGGCTTTGGAAACAAATTTCTATTATTACCCCAATGCAATTTTCCACT 1260
Qy 1261 TTGCCATCCCTAATAAGCTGGAACCTATCGTCCAACTCTCTGTGCTTTTCCCTATAACT 1320
Db 1261 TTGCCATCCCTAATAAGCTGGAACCTATCGTCCAACTCTCTGTGCTTTTCCCTATAACT 1320
Qy 1321 GGGTTACATGTTTAACTCAGCTTAAATTAACAGGAATCATGCTTACAGAGCTTTGATA 1380
Db 1321 GGGTTACATGTTTAACTCAGCTTAAATTAACAGGAATCATGCTTACAGAGCTTTGATA 1380
Qy 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGTGT 1440
Db 1381 TCATCTGAAACTTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAAGTGTGT 1440
Qy 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
Db 1441 GCATTTGGAGTGTGAGAAATGCTTAAAGATTTCTAATCAATGGAATAAAGGTGACAAAC 1500
Qy 1501 AGCAGTATGGAAGACCTTTTATAAGAAATGCTTGAATGTTTTCAGGCTCAAGATGAAAGT 1560
Db 1501 AGCAGTATGGAAGACCTTTTATAAGAAATGCTTGAATGTTTTCAGGCTCAAGATGAAAGT 1560
Qy 1561 GACCTTGAAGATTTTCTGCTTGTGAGGAAAGCTTGAAGGCTTCAATTCAGTGCAG 1620
Db 1561 GACCTTGAAGATTTTCTGCTTGTGAGGAAAGCTTGAAGGCTTCAATTCAGTGCAG 1620
Qy 1621 TGTTCACTTCCCAAGGCCCTTCAAACTTGTGAACACCTGCTTGTGATGGCTGGCTGATC 1680
Db 1621 TGTTCACTTCCCAAGGCCCTTCAAACTTGTGAACACCTGCTTGTGATGGCTGGCTGATC 1680
Qy 1681 AGAATGGAGTGTGAGCAATAGCAGTCTGCACTTCTGCTTAATGCTTTGGTGAATTTCA 1740
Db 1681 AGAATGGAGTGTGAGCAATAGCAGTCTGCACTTCTGCTTAATGCTTTGGTGAATTTCA 1740
Qy 1741 ACAGTTTTTCAGATCCCTCTGTACATTTTCCCAATTAACCTGTTAAATTTGGGGTCAATCGA 1800
Db 1741 ACAGTTTTTCAGATCCCTCTGTACATTTTCCCAATTAACCTGTTAAATTTGGGGTCAATCGA 1800
Qy 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCGTGGCTGGTGGTGGATGGCTTCACT 1860
Db 1801 GCAGTGAACATGCTCAAGGAGTCTCCAGTCCCGTGGCTGGTGGTGGATGGCTTCACT 1860
Qy 1861 TTTGGCAGTTTGCACGACATGGTCCCTGGTGGGAGAAATGGGTTGGTGGCATGTCATT 1920
Db 1861 TTTGGCAGTTTGCACGACATGGTCCCTGGTGGGAGAAATGGGTTGGTGGCATGTCATT 1920
Qy 1921 GGTTTTTTGTCCATTTTTTGTTCAGAAATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Db 1921 GGTTTTTTGTCCATTTTTTGTTCAGAAATCATCTGTTTTTCTGCTTACTCTGCGAGCCCTG 1980
Qy 1981 GAGCGTGGGTTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Db 1981 GAGCGTGGGTTCTCTGTGAAATATTTGCAAAATTTGAAACGAAAGCTCCATTTTCTAGC 2040
Qy 2041 CTGAAAGTAAATCATTTTTTGTCTCTGTGCCCTGTGCTTGAACATGGCCGAGTTCCCTCTG 2100
Db 2041 CTGAAAGTAAATCATTTTTTGTCTCTGTGCCCTGTGCTTGAACATGGCCGAGTTCCCTCTG 2100
Qy 2101 CTGGTGGCAGCAAGATATGGCGCTCCCTCTGCTGCTGCTTTGGCTTTTGGGAGGCC 2160
Db 2101 CTGGTGGCAGCAAGATATGGCGCTCCCTCTGCTGCTGCTTTGGCTTTTGGGAGGCC 2160
Qy 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220
Db 2161 AGCACCATGGGCTACATGGTGGCTCTCATCTTGTCTCAATTCCTTGTCTTCTCATGATG 2220

Adams, M.D. and Cargill, M. Direct Submission Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA This sequence was made by sequencing genomic exons and ordering them based on alignment FEATURES source Location/Qualifiers 1..2593 /organism="Pan troglodytes" /mol_type="genomic DNA" /db_xref="taxon:9598" <1..>2593 /gene="GPR49" /locus_tag="HCM4322"									
ORIGIN									
Query Match 71.9%; Score 1958; DB 9; Length 2593; Best Local Similarity 75.7%; Pred. No. 0; Matches 1964; Conservative 0; Mismatches 629; Indels 0; Gaps 0;									
QY	132	CAGGATGTTGCTCAGGGTGGACTGCTCCGACCTGGGGCTCTCGAGCTGCCTTCCAACT	191						
DB	1	CAGGATGTTGCTCAGGGTNNNNNNNTCCGACNNNGGCTCTCGAGCTGCCTTCCAACT	60						
QY	192	CAGCGCTTCACTCTCACTAGACCTCAGTATGAACAACATCAGTCAGCTGCCGAA	251						
DB	61	CAGCGCTTCACTCTCACTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	120						
QY	252	TCCCTGCCCCAGTCTCCGCTTCTCTGGAGGAGTTACGTCTTGGGGAAACGCTCTGACATA	311						
DB	121	NN	180						
QY	312	CATTCCTCCAGGGAGCATTCACCTGGCCCTTTACAGTCTTAAAGTCTTATGCTGCAGATAA	371						
DB	181	NN	240						
QY	372	TCAGCTAAGACAGTACCACAGAGCTCTGCAGAAATTTGGAGCGCTTCAATCCCTGGC	431						
DB	241	NN	300						
QY	432	TCGTGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGCGTGCATTCCT	491						
DB	301	TCGTGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGTTTCAGTGGCGTGCATTCCT	360						
QY	492	GAGGCACCTGTGGCTGGATGACAATCGTTTAAACAGAAATCCCGCTCCAGGCTTTTGAAG	551						
DB	361	GAGGCACCTGTGGCTGGATGACAATCGTTTAAACAGAAATCCCGCTTTTGAAG	420						
QY	552	TTTATCGGCAATGCAAGCCATGACCTTGGCCCTGAAACAAAATAACACATACCAACTA	611						
DB	421	TTTATCGGCAATGCAAGCCATGACCTTGGCCCTGAAACAAAATAACACATACCAACTA	480						
QY	612	TGCGCTTTGGAACCTCTCCAGCTTGTAGTCTTACATCTCCATAACATAGAAATCCACTC	671						
DB	481	TGCGCTTTGGAACCTCTCCAGCTTGTAGTCTTACATCTCCATAACATAGAAATCCACTC	540						
QY	672	CCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGATTTAAATTA	731						
DB	541	CCTGGGAAAGAAATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTINNNNNNNNNNN	600						
QY	732	TAACTTTGATGAAATTTCCCACTGCAATTAGGACACTCTCCAACTTTAAGAACTAGGATT	791						
DB	601	NN	660						
QY	792	TCATAGCAATATCAGTCCATACCTCAGAAAGCATTTGTAGGCAACCTTCTCTTAT	851						
DB	661	TCATAGCAATATCAGTCCATACCTCAGAAAGCATTTGTAGGCAACCTTCTCTTAT	720						
QY	852	TACAAATACATTTCTATGACAATCCCACTCCAAATTTGTGGAGATCTGCTTTTCAACATTT	911						
DB	721	TACAATNCAATTTCTATGACAATCCCACTCCAAATTTGTGGAGATCTGCTTTTCAACATTT	780						
QY	912	ACCTGAACTAAGAACACTGACTCTCTGAATGGTGGCTTCAAAATAACTGAATTTCTGATT	971						

DB	781	ACCTGAACTNN	840						
QY	972	AACTGGAACTGCAAACTGGAGAGCTGACTTTAACTGGAGCACAGATCTCATCTCTTCC	1031						
DB	841	AACTGGAACTNN	900						
QY	1032	TCAAAACCGTCTGCAATCAGTTACCTAACTCCAAAGTGTAGATCTGTCTTACAACTATT	1091						
DB	901	NN	960						
QY	1092	AGAAGATTATCCAGCTTTTTCAGTCTGCCAAAGCTTCCAGAAATTTGACCTAAGACATAA	1151						
DB	961	AGAAGATTATCCAGCTTTTTCAGTCTGCCAAAGCTTCCAGAAATTTGACCTAAGACATAA	1020						
QY	1152	TGAAATCTACGAAATTTAAAGTTGACACTTTTCCAGCAGTGTGCTTAGCCTCCGCTGAA	1211						
DB	1021	TGAAATCTNN	1080						
QY	1212	TTTGGCTTGGAAACAAAATTTGCTATTATTCACCCCAATGCAATTTTCCACTTTCCTCT	1271						
DB	1081	NN	1140						
QY	1272	AATAAGCTGGACCTATCTGCCAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1331						
DB	1141	NN	1200						
QY	1332	TTTAACTCACTTAAATTTAAACAGGAAATCATGCCCTTACAGAGCTTGATATCATCTCAAAA	1391						
DB	1201	TTTAACTCACTTAAATTTAAACAGGAAATCATGCCCTTACAGAGCTTGATATCATCTCAAAA	1260						
QY	1392	CTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGTGCAATTTGGAGT	1451						
DB	1261	CTTTCAGAACTCAAGGTTATAGAAATGCTTATGCTTACAGTGTGTGCAATTTGGAGT	1320						
QY	1452	GTGTGAGATGCTATAGATTTCTTAATCAATGGAATGAAGCTGACCAAGCAGATATGA	1511						
DB	1321	GTGTGAGATGCTATAGATTTCTTAATCAATGGAATGAAGCTGACCAAGCAGATATGA	1380						
QY	1512	CGACCTTCATAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACCTGCACTTGAAGA	1571						
DB	1381	CGNNCTTCATAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAACCTGCACTTGAAGA	1440						
QY	1572	TTTCTGTGTGCTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGTGTGCACTTTCACCTTC	1631						
DB	1441	TTTCTGTGTGCTTTGAGGAAGACCTGAAAGCCCTTCAATCAGTGTGTGCACTTTCACCTTC	1500						
QY	1632	CCAGGCCCTTCAAAACCTGTGAACACTGCTTGTGATGGCTGGCTGATCAGAAATTTGGAGT	1691						
DB	1501	CCAGNN	1560						
QY	1692	GTGGACCATAGCAGTCTTGGCACTTACTTGTAACTGTTGGTGACTTCAACAGTTTTCAG	1751						
DB	1561	NN	1620						
QY	1752	ATPCCCTCTGTACATTTCCCCCATTTAAACTGTTAAATGGGGTTCATCGCAGCAGTGAACAT	1811						
DB	1621	NN	1680						
QY	1812	GCTCAGGGAGTCTCAGTGGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1871						
DB	1681	GCTCAGGGAGNNNNNAGTGGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	1740						
QY	1872	TGCAGACATGTTGCTGCTGGTGGGAGATGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1931						
DB	1741	TGCAGACATGTTGCTGCTGGTGGGAGATGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	1800						
QY	1932	CATTTTGTCTTCAGAAATCATCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1991						
DB	1801	CATTTTGTCTTCAGAAATCATCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1860						
QY	1992	CTCTGTGAAATATTTCTGCAAAATTTGAAACGAAAGCTTCAATTTTCTAGCTGCAAGATAT	2051						

QY 1718 CTTGTAATGCTTTGGTGAATCAACAGATTTTTCAGATCCCTCTGTACATTTCCCCCATTA 1777
Db 721 CTTGTAATGCTTTGGTGAATCAACAGATTTTTCAGATCCCTCTGTACATTTCCCCCATTA 780
QY 1778 AACTGTTAAATGGGGTTCATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGC 1837
Db 781 AACTGTTAAATGGGGTTCATCGCAGCAGTGAACATGCTCACGGGAGTCTCCAGTGCCTGC 840
QY 1838 TGGCTGGTGTGATGCGTTCACATTTTGGCAGCTTTTGGCAGCATGCTCGCTGGTGGGAGA 1897
Db 841 TGGCTGGTGTGATGCGTTCACATTTTGGCAGCTTTTGGCAGCATGCTCGCTGGTGGGAGA 900
QY 1898 ATGGGGTGGTGGTGGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1957
Db 901 ATGGGGTGGTGGTGGTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
QY 1958 TCCTGCTTACTCTGCGAGCCCTGGAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTG 2017
Db 961 TCCGCTACTCTGCGAGCCCTGGAGCGTGGGTCTCTGTGAAATATTTCTGCAAAATTTG 1020
QY 2018 AAACGAAAGCTCCATTTTCTAGCTGGAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCT 2077
Db 1021 AAACGAAAGCTCCATTTTCTAGCTGGAAGTAATCATTTTGTCTGCTGCTGCTGCTGCTGCT 1079
QY 2078 TGACCATGGCGCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 1080 GACCHGCGCCGTTTCCCGCCCGG 1102

RESULT 5

AL541959
DEFINITION AL541959 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE007YC19
5-PRIME, mRNA sequence.
ACCESSION AL541959
VERSION AL541959.3 GI:45717535
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1060)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On Feb 15, 2001 this sequence version replaced gi:30546637.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: segre@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 7021.f
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?g=CS0DE007AB10QPI&c=7021.f.

FEATURES

source
1..1060
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE007YC19"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN

Query Match 35.9%; Score 979.2; DB 1; Length 1060;
Best Local Similarity 98.3%; Pred. No. 5.4e-273;
Matches 993; Conservative 6; Mismatches 9; Indels 2; Gaps 1;
QY 683 AATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGATTTAAATTTACAAATACCTTTGATG 742
Db 1 AATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGATTTAAATTTACAAATACCTTTGATG 60
QY 743 AATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGATTTAAATTTACAAATACCTTTGATG 802
Db 61 AATGCTTTGATGGGCTCCACAGCCTTAGAGACTTTAGATTTAAATTTACAAATACCTTTGATG 120
QY 803 ATATCAGCTCCATACCTGAGAAAGCATTTGTAGGCAACCTCTCTCTTATTAACAATACATTT 862
Db 121 ATATCAGCTCCATACCTGAGAAAGCATTTGTAGGCAACCTCTCTCTTATTAACAATACATTT 180
QY 863 TCTATGACAAATCCCAATCCCAATTTGTGGAGATCTGCTTTTCAACATTTTACCTGAACTAA 922
Db 181 TCTATGACAAATCCCAATCCCAATTTGTGGAGATCTGCTTTTCAACATTTTACCTGAACTAA 240
QY 923 GAAACATGACTCTGAAATGGTGCCTCAAAATTAACCTGAATTTCTGATTTAACTGGAATG 982
Db 241 GAAACATGACTCTGAAATGGTGCCTCAAAATTAACCTGAATTTCTGATTTAACTGGAATG 300
QY 983 CAACCTGGAGCTGACTTTAACTGAGCAGACAGATCTCATCTCTTCTCTCAACCGCTCT 1042
Db 301 CAACCTGGAGCTGACTTTAACTGAGCAGACAGATCTCATCTCTTCTCTCAACCGCTCT 360
QY 1043 GCAATCAGTTACCTAACTCTCCAAGTGTAGATCTGTCTTACAACTATTTAGAAAGATTTAC 1102
Db 361 GCAATCAGTTACCTAACTCTCCAAGTGTAGATCTGTCTTACAACTATTTAGAAAGATTTAC 420
QY 1103 CCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTTGACCTAAGACATATTAATGAAATCTAG 1162
Db 421 CCAGTTTTTTCAGTCTGCCAAAAGCTTCAGAAAAATTTGACCTAAGACATATTAATGAAATCTAG 480
QY 1163 AAATTAAGTTGACACATTTCCAGCAGTTGCTTAGCTCCGATCGTGATTTGGCTTGA 1222
Db 481 AAATTAAGTTGACACATTTCCAGCAGTTGCTTAGCTCCGATCGTGATTTGGCTTGA 540
QY 1223 ACAAATTTGCTATTATTTCACCCCAATGCTTTTCCACATTTTCCATCTCCATTAATAAGCTGG 1282
Db 541 ACAAATTTGCTATTATTTCACCCCAATGCTTTTCCACATTTTCCATCTCCATTAATAAGCTGG 600
QY 1283 ACCTATCGTCCAACTCTCTGCTCTTTTCTTAACTAGGTTACATGTTTAACTCACT 1342
Db 601 ACCTATCGTCCAACTCTCTGCTCTTTTCTTAACTAGGTTACATGTTTAACTCACT 660
QY 1343 TAAATTAACAGGAAATCATGCTTACAGACTTGTATCATCTGAAAACTTTCCAGAAC 1402
Db 661 TAAATTAACAGGAAATCATGCTTACAGACTTGTATCATCTGAAAACTTTCCAGAAC 720
QY 1403 TCAAGGTTTATAGAATGCTTATGCTTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1462
Db 721 TCAAGGTTTATAGAATGCTTATGCTTACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 1463 CCTATAGATTTTCAATCAATGGAATTAAGGTGACACAGCAGTATGACGACCTTTCATA 1522
Db 781 CCTATAGATTTTCAATCAATGGAATTAAGGTGACACAGCAGTATGACGACCTTTCATA 840
QY 1523 AGAAAGTGTGGAATGTTTCAGGCTCAAGATGAACCTGAGCTTGAAGATTTCTGCTTG 1582
Db 841 AGAAAGTGTGGAATGTTTCAGGCTCAAGATGAACCTGAGCTTGAAGATTTCTGCTTG 900
QY 1583 ACTTTGGAGAGACTGAAAGCCCTTCAATTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1642
Db 901 ACTTTGGAGAGACTGAAAGCCCTTCAATTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 958
QY 1643 TCAAAACCTGTGGAACACTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1692
Db 959 TCAAAACCTGTGGAACACTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008

RESULT 6	AK047873	4796 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK047873				
DEFINITION	Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone:CL130018C02 product:G protein-coupled receptor 49, full insert sequence.				
ACCESSION	AK047873				
VERSION	AK047873.1	GI:26339017			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 93279253 10349636				
REFERENCE	2 Carninci, P., Shibata, Y., Hayata, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) 20499374 11042159				
REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagao, K., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitesunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Ikegami, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Nishigaki, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 11076861				
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
REFERENCE	5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 4796)				
REFERENCE	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, M., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsubara, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)				
TITLE	Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suhiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)				
JOURNAL	CDNA library was prepared and sequenced in Mouse Genome				
COMMENT					

Query Match	29.6%;	Score	805.8;	DB	3;	Length	3583;
Best Local Similarity	59.8%;	Pred.	No. 2.9e-222;				
Matches 1503; Conservative	0;	Mismatches	982;	Indels	28;	Gaps	8;

	QY	14	GGCTCGGTGTCCTCCTGCTTGCCCTGCTGCTGTCGACAGTGGCGACCGGGGGGAGTCTCTC	73
Db	111	GGCTTCCTGGCGCTGTGGCTTTGGCTGTGCTGTGGCATCGGC CGCGCGGGCAGCGACC	170	
QY	74	CCAGGTCTGGTGTGTTGCTGAGGGGCTGCCCCACACACTGTCTCAATGCGAGCCCGCAGCGCA	133	
Db	171	CCCAGCGTGGCCCGGGGCGTCCCGCTGCCCGCTCCCTGCCACTGCTCCAGAGAGCGGCA	230	
QY	134	GGATGTGCTCAGGTTGGAATGCTCCCGACTGGGGCTCTCGAGCTGCTTCCAACCTCA	193	
Db	231	TCAATGCTGTC---GCTGACTGCTCGAGCTCGGGCTCTCAGTGTGTGCTCTGCGGAGCTGG	287	
QY	194	GGCTCTTCACTCTCTACTAGACTCTCAGTATGAACAACATCAGTCACTGCTCTCCGAATC	253	
Db	288	ACCCCTGACGGCTTACCTTAGAAGCTCAGTATGAACAACCTCACGGAGCTTCAGCGGGTC	347	
QY	254	CCCTGCCAGTCTCGGCTTCTCGGAGAGTTACGTCTTGGGGAAACGCTCTGACATACA	313	
Db	348	TCATTCCACCACCTCGCGTTCTCGAGGAGCTGGGGCTCTCAGGGAACCACTCTCTCACACA	407	
QY	314	TTCCCAAGGAGGATTCACATGGCTTTTACAGTCTTAAGAATCTTATGCTSCAGAAATAATC	373	
Db	408	TCCGGGACAGGCAATCTCTGGCTCTCAGAGCTCAAATTTAATGCTGACAGCAACC	467	
QY	374	AGCTAAGACACGTACCCACAGAAGCTCTGCAGAAATTTGCGAAGCCTTCAATCCCTGGCTC	433	
Db	468	AGCTCCGTGGGATCCACAGCAGAGAGCACTATGGGAGCTGCCAGCTGCAGTCTGCTGCGCC	527	
QY	434	TGATGCTAACCATCATCAGCTATGTGCCCCCAAGCTGTTTTCAGTGGCGCTGCATTCCTGTA	493	
Db	528	TAGATGCTAATTCATCTCCCTGGTCCCTGAGAGAAGCTTTGAGGGGCTCTCTCTCCCTCC	587	
QY	494	GGCACCTGTGGCTGGATGACAATGCGTTAAACAGAAATCCCCGCTCCAGGCTTTTAGAAGTT	553	
Db	588	GCCACCTTGCTGGATGACAATGCACTACCGAGATCCCGTTCAGAGCTCTCAACAACC	647	
QY	554	TATCGGCAITGCAAGCATGACCTTGGCCCTGAAACAAATACACCAATACCAAGACTATG	613	
Db	648	TTCTCGCTTACAAAGCCATGACCTTGGCTCTCAACCATATCCGCCACATCCCTGACTATG	707	
QY	614	CCTTTGGAAACCTCTCCAGCTTGGTACTCTACATCTCCATAACATAGAAATCCACTCCC	673	
Db	708	CCTTTCAGAAACCTCACCAAGCTTTGGTGTGTCATCTACATAACACCGCATTCAGCATG	767	
QY	674	TGGGAAAGAAATGCTTTGATGGGCTCCACAGCCCTAGAGACTTTAGATTTTAAATTCACAATA	733	
Db	768	TGGGGACCCACAGCTTCAGGGGGCTGCACAACTTGGAGACACTAGACCTGAACTATAATG	827	
QY	734	ACCTTGATGAATTTCCCACTGCAATTAGGACACTCTCCAACCTTAAAGAACTAGGATTTTC	793	
Db	828	AGTGTGAGGAGTTTCCCTTGGCTATCCGAGCCCTGGCGAGGCTGCAGGAAATGGGTTTCC	887	
QY	794	ATAGCAACAATATCAGGTTCGATACCTGAGAAAAGCATTTGTAGGCAACCCCTTCTCTATTAT	853	
Db	888	ATAACAACAATCAAGGCTATCCAGAGAAAGCCTTCATGGGCGAGCCCTCTCTCTGCAGA	947	
QY	854	CAATACATTTCTATGACAAATCCCAATCOAATTTGTTGGGAGATCTGCTTTTCAACATTTAC	913	
Db	948	CAATACATTTTATGACAACCAATCCAGTTTGTGGGAAGGTGAGCATTTCCAGTACCTGT	1007	
QY	914	CTGAACATAAGAACACTGACTCTGAATGGTGCCTCACAAATAACTGAATTTCTTGATTAA	973	
Db	1008	CTAACTGCAATACGTAATCTTTGAATGGTGCACATGATATCCAGAGAGTTCCGACAGCTCA	1067	
QY	974	CTGGAAGCTCAAAACCTGGAGAGTCTGACTTTAACTGGAGCACAGATCTCACTCTTCCCTC	1033	
Db	1068	AAGGACCACTAGCTCGGAGATCTGACCTGACCCGCTGGGGGATCAGACTGCTGCCAC	1127	

Qy	1034	AAACCGTCTGCAATCAGTTTACCTAAATCTCCAAAGTGTAGATCTGTCTTTCAACACCTATTAG	1099
Db	1128	CGGGAGTGTGCCAACACGTGCTAGGCTCCGAATCTCTGGAGCTGTCTCAATAATFCAGATCG	1187
Qy	1094	AAGATTATCCACCAAGTTTTCAG-GTCTGCCAAAAGCTTTCAGAAAATTCAGCTTAAGACATAAT	1152
Db	1188	AGAGTATTACCCAGCTTGCGCACAGGTGTACAGACTGGAGAAATTTGGCTTCCGACATTAAC	1247
Qy	1153	GAATCTTACGAAATATAAAGTTTGACACTTTTCCAGCAGTTGTCTAGCCTCCGATCGCTGAAT	1212
Db	1248	CGGATCAAGGAATTTGGTGCAGATACCTTCAGCCAGCTGGCTCCTTGCAGACTTTAGAC	1307
Qy	1213	TTGGCTTGGAAACAAAATTTGCTATTATTATCACCCCAATGCAATTTTTCACATTTTGCATCCCTA	1272
Db	1308	CTGAGTTGGAAATGCGCATCCGTGCGCATCCACCCCTGAGGCTTTCTCAACCCCTTCGATCCTTG	1367
Qy	1273	ATAAGCTTGACCTATCGTCCAACTTCTGTCTGTCTTTTCCATAACTGGGTTACATGGT	1332
Db	1368	GTTAAGCTGACCTGACTGACAAACCAAGCTGACCACACTGCCCTGTGCGCTGGGCTGGAGGC	1427
Qy	1333	TTAACTCACCTTAAAAATTAAACAGGAATCATGCCCTTACAGAGCTTGATATCATCTGAAAAAC	1392
Db	1428	CTGATGCACCTGAAGCTCAAGGGAATTTGGCCCTGTCTCAGGCTTCTCCAAGGACAGT	1487
Qy	1393	TTTTCCAGAACTCAAGGTTATPAGAAATGCCCTTATGCTTTACAGTGTGTGCAATTTGAGTG	1452
Db	1488	TTCCAAAACTGAGGATCCTGGAGGTGCCCTACGCCCTACCAAGTGTGTGCTACCGCATC	1547
Qy	1453	TGTGAGNATGCCCTATAAGATTTCTAATCAATGGAATAAAGGTGAC---AACAGCAGTATG	1509
Db	1548	TGTGCCAGCTTCTTCAAGACCTCTGGGCACTTGGGCACTGGCGCCGAGGACTTTTATCCAGAGGAA	1607
Qy	1510	GACGACCTTCATAAAGAAAGATGCTGGAAATGTTTTCAGGCTCAAGATGAAAGCTGTGACCTTGAA	1569
Db	1608	GAGAGGCCAATAAGAGGCCCTTGGGTCTCCTTGTGTGGAACAGCTGAGAACCACTATGAC	1667
Qy	1570	GATTTCTGCTGTGACTTTTGAG-----GAAGACCTGAAAAGCCCTTCATTCAAGTGCAAG	1620
Db	1668	CTAGACCTTGATGAGCTCCAGATGGGACAGAGGACTCAAAAGCCACACCCAGTGTCCAG	1727
Qy	1621	TGTTACCTTCCCGAGGCCCTTCAA-ACCTGTGAACACCTGCTGTGATGGCTGGCTGAT	1679
Db	1728	TGCAGCCCTGTTCCAGGCCCTTCAAGGCCCTCGAGCACCTCTTTGAGAGCTTGGGGCAT	1787
Qy	1680	CAGAAATGGAGTGTGAGCAATACAGTTCCTGGCACTTACTTGTAAATGCTTGTGTGACTTC	1739
Db	1788	CCGCTTGTGTGTGGGCCATCGTGTACTCTCCGTACTCTGTGAACGGCTGTGTGCTGCT	1847
Qy	1740	AACAGTTTT---CAGATCCCTCTGTACATFTTCCCCCAATTAACACTGTTAATTTGGGGTCAT	1796
Db	1848	GACAGTCTTTGCCAGCGGACCCAGCCCGTGTCCCCCGTCAAGCTTGTGTGGTGGTCCGAT	1907
Qy	1797	CGCAGCAGTGAAATGCTCTACGGGAGTCTCCAGTCCGTGTGCTGGCTGGTGTGATCGGTT	1856
Db	1908	GGCAGGCGCCAAAGCCCTGTGCGGCAATTTCTGTGGTCTCTCGCCCTCGGTGACAGCCCTT	1967
Qy	1857	CACCTTTGGCAGCTTTTGACGACATGGTGCTGGTGGGAGAAATGGGGTGTGGTTGCCATGT	1916
Db	1968	GACCTATGGTCACTTGTGCTGAGTATGAGAGCCCTCGGAGAGCGGCTGTGGCTGCCAGGC	2027
Qy	1917	CATTGGTTTTTTTGTCAATTTTTGCTTCAGAAATCATCTGTTTTTCCGTCTTACTCTGGCAGC	1976
Db	2028	TACGGGCTTCTGGCTGTCTGGGTTTCAGAGGCTCGGTGCTGCTGCTCACACTGGCGGC	2087
Qy	1977	CCTGGAGCGTGGGTCTCTGTGAAAT-ATTCTGCAAAATTTGAAACGAAAGCTCCATTTT	2035
Db	2088	CGTGCAGTGCAGCATCTCCGTGACCTCGTCCGATCCAGAGCTACGGGAAGCGCGCTCGCCTG	2147
Qy	2036	CTAGCTGAAAGTAATCATTTTTGCTCTGTGCCCTGTGGCTTGAACCATGGCCGACAGTTTC	2095
Db	2148	GCACGCTCCGCGCAGGCGCATTTGGAATGCTGCGCTGGCCGGGCTGGCCGACAGCACTGC	2207
Qy	2096	CCCTGCTGGGCTGGCAGCAAGTATGGGGCCCTCCCTCTCTGCTCTGCCCTTTTGGGCTTTGGGG	2155

QY	1034	AAACCGTCTGCAATCAGTTACCTTAATCTCCAAGTCTAGATCTGTCTTCAACCTATTAG	1093
Db	1128	CGGGAGTGTGCCAACAGCTGCCTAGGTCCTGAATCTCGGAGTGTCTCATTAATCAGATCG	1187

QY 1094 AGATTTACCCAGTTTTTCA-GTCTGCCAAAGCTTCAGAAAATTGACCTTAAGACATAAT 1152

Db	1188	AGGAGTTATCCCAAGCCTGGCACAGGTTGTGAGAAGCTTGGAGGAATTTGGCTCCGACATAAC	1244
Qy	1153	GAAATCTACGAATAATAAGTTGCACATTTCCAGCAGTTGCTTTAGCCTCCGATCCGCTGAAT	1212

Db 1248 CGGATCAAGGAATTGGTGCAGATACCTTCAGCCAGCTGGGTCCTTTCGAAGCTTTAGAC 1307

Qy 1213 TTGGCTTGGAAACAAAATTTGCTATTATTCACCCCAATTCGATTTCCACTTTGGCATCCCTTA 1277

Dh 1308 CTGAGTTGGAATGCCATCCGTCGCCATCCACCCCTGAGGCTTTCTCAACCCCTTCGATCCTTG 1367

[illegible]

DB	QY
1333	TTAACTCACTTAAAAATTAAACAGGAAATCATGCGCTTACAGAGCTTGATATCATCTGAAAAC
1368	GTAAAGCTGGAGCCCTGATCTGATGACAAACACAGCTGACACACATGCTCCCTGGCTGGGCTGGGAGGC
1421	

Db 1428 CTGATGCACCTGAAGCTCAAAAGGAACTTGGCCCTGTCTCAGGCCCTCTCCAAGGACAGT 1487

Ov 1393 TTTCCAGAACTCAAGGTTATAGAAATGCCCTTATGCTTACCAAGTGCCTGTGCATTTGGAGTG 1452

1488 TTCCAAAACCTGAGGATCCTGGAGGTGCCCTACGCCCTACCAAGTCTGTGCTTACGGCATC 1547

		TGAGGAATGCCTTATAGATTCTTAATCAATGGNATAAAGGTGAC---	AACAGCAGCATATG	1503
Qy	1453			
			TGGCCAGCTTCCTTCAAGACTCTTGGCAGCTTGGCAGCGAGGACTTTTCATCCAGAGGAA	1607
Dδ	1548			

QY 1510 GAGGACCTTCATAAGAAAGATGCTGGAAATGTTTCAGGCTCAAGATGAAACGTGACCTTGAA 1569

DB	1570	GATTTCTGCTTGACTTTGAG-----GAAGACCTGAAAGCCCTTCATTTCAAGTGAC	1620
DB	1608	GAGGAGGACCAAGAGAGGCCCCCTGGGTCTCCTTGCTTGACAAAGCTGAGAAACCACTATGAC	1668

[illegible]

Qy	1624	1625	1626	1627	1628	1629	1630	1631	1632	1633	1634	1635	1636	1637	1638	1639	1640	1641	1642	1643	1644	1645	1646	1647	1648	1649	1650	1651	1652	1653	1654	1655	1656	1657	1658	1659	1660	1661	1662	1663	1664	1665	1666	1667	1668	1669	1670	1671	1672	1673	1674	1675	1676	1677	1678	1679	1680	1681	1682	1683	1684	1685	1686	1687	1688	1689	1690	1691	1692	1693	1694	1695	1696	1697	1698	1699	1700	1701	1702	1703	1704	1705	1706	1707	1708	1709	1710	1711	1712	1713	1714	1715	1716	1717	1718	1719	1720	1721	1722	1723	1724	1725	1726	1727	1728	1729	1730	1731	1732	1733	1734	1735	1736	1737	1738	1739	1740	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759	1760	1761	1762	1763	1764	1765	1766	1767	1768	1769	1770	1771	1772	1773	1774	1775	1776	1777	1778	1779	1780	1781	1782	1783	1784	1785	1786	1787	1788	1789	1790	1791	1792	1793	1794	1795	1796	1797	1798	1799	1800
Db	1624	1625	1626	1627	1628	1629	1630	1631	1632	1633	1634	1635	1636	1637	1638	1639	1640	1641	1642	1643	1644	1645	1646	1647	1648	1649	1650	1651	1652	1653	1654	1655	1656	1657	1658	1659	1660	1661	1662	1663	1664	1665	1666	1667	1668	1669	1670	1671	1672	1673	1674	1675	1676	1677	1678	1679	1680	1681	1682	1683	1684	1685	1686	1687	1688	1689	1690	1691	1692	1693	1694	1695	1696	1697	1698	1699	1700	1701	1702	1703	1704	1705	1706	1707	1708	1709	1710	1711	1712	1713	1714	1715	1716	1717	1718	1719	1720	1721	1722	1723	1724	1725	1726	1727	1728	1729	1730	1731	1732	1733	1734	1735	1736	1737	1738	1739	1740	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759	1760	1761	1762	1763	1764	1765	1766	1767	1768	1769	1770	1771	1772	1773	1774	1775	1776	1777	1778	1779	1780	1781	1782	1783	1784	1785	1786	1787	1788	1789	1790	1791	1792	1793	1794	1795	1796	1797	1798	1799	1800

1680 CAGAAATTGGAGTGTGGACCAATAGCAGTTCTTGGCACTTACTTCTGTAAATGCTTTTGGTGACTTC 1739

Db CGCCCTTCTGTGTGGGCGCATCGTGTTACTCTCCGTACTCTGTAAACGGGCTGGTGTGCT 1847

QY 1740 AACAGTTTT---CAGATCCCCCTCTGTACATTTCCTCCCATTAACCTGTAATTGGGGTCAT 1796

Accession	Gene	Accession	Gene
1848	GACAGTCTTTTGCCAGCGGACCCAGCCGCTGCTCCCGCTCAAGCTTTGTGGTGGGTGGCAT	1907	CGCAGTCTTTTGCCAGCGGACCCAGCCGCTGCTCCCGCTCAAGCTTTGTGGTGGGTGGCAT
Db			
Qy	1797	CGCAGTCTTTTGCCAGCGGACCCAGCCGCTGCTCCCGCTCAAGCTTTGTGGTGGGTGGCAT	1856

Db 1908 GG CAG GCG CCA AG CCG CCG TGT CCG GCA TTT C CTG TGT CT C TGG CCG TCG TCG GCG CCG TT 1967

QY 1857 CACTTTGGCAGCTTTGACGACAATGTCCTGGTGGGAGAAATGGGTGTGTTGCCAATG 1911

Dh 1968 GACCTATGTTCAGTTCGCTGAGTATGGAGCCGCTGGGAGAGCGGTCTGGGTCTGCCAGGC 2077

QY	1917	CA	TG	GG	TTTT	TG	TC	CA	AT	TTTT	TG	CT	TC	CA	GA	AT	CA	TG	TTTT	TC	CT	TT	TC	CT	TC	GG	CAGC	1976			
ph	2028	T	A	C	G	G	G	T	T	C	C	G	G	T	T	C	C	G	G	T	T	C	C	G	G	T	T	C	C	G	G

2026
1977
QY

Db	2088	CGTGCAATGCAGCATCTCCGTACCTCGAGCTACGGGAAGSCCGCTGCGCTG	2147
QV	2036	CTAGCCTGAAGTAATCATTTTGCTGTGCGCTGTGCGCTTGACCATGGCCGCAATTG	2095

2148 GCAGCGTCCGCGCAGGCGCACTTGGATGCTCCCGCTGGCCGGGCTGGCCGCGAGCACTGC 2207

QY 2036 CCCTGCTGGTGACAGCAAGTATGGCGCCTCCCCCTCTCTGCTGCTGCCCTTGGCCCTTTTGGGG 2155

Db	2208	CGCTGGCGCTCGGTGGGAGAGTAATGGCGCCTCCCCACACTCTGCTGCGCCTACGCCCCACCCG	2267
Qy	2156	AGCCACAG-----CACCATGGGGCTACATGTCGCTCTCATCTTGTCTCAATTCCTTTTGC	2208
Db	2268	AGGCGCGCGCGCTGCGCTTGGCTGTAGCCCTGGTGATGATGAACCTCGCTCTGC	2327
Qy	2209	TTCTCATGATGACCATTTGCTACACCAAGCTCTACTGCAATTTGGACAGGAGACCTG	2268
Db	2328	TTCTGTGTGTGGCGCGGCTCATCAAGCTCTACTGTACCTTGCACCGGGGTGACTTT	2387
Qy	2269	GAGAAATTTGGGACTCTCTATGGTAAAAACAATTCGCCCTGTGTCTTTCACCAACTGC	2328
Db	2388	GAGCCGTGTGGGACTCGCGCCATGGTGGCGCACGTGSCCTGCCTCATCTTTGCAGATGC	2447
Qy	2329	ATCGTAAACGCCCTGTGGCTTTCTTGTCCTTCTCTCTTTTAAATAAACCTTACATTTATC	2388
Db	2448	CTCCTCTACTGCCCCGTGGCTTCTCTCAGCTTTTGGCTCTATGCTGGGCGCTCTTCCCTGTC	2507
Qy	2389	AGTCTCGAAGTAATTAAGTTTATCTTCTGGTGGTAGTCCCACTTCTCGTCATGCTCAAT	2448
Db	2508	ACCCCGAGGCTGTCAAGTCAGTCTTCTGGTGGTGTGCTCTGTGCTGTGCTGCTCAAC	2567
Qy	2449	CCGCTTCTTACATCTTGTTCATTCCTCACTTTAAAGGAGGATCTGGTNGAGCCT	2501
Db	2568	CCACTGTCTACTGTCTTTAAACCCCTCACTTCGGGATGACCTTGGCGGCT	2620

RESULT	8
CK449409	
LOCUS	CK449409 940 bp mRNA linear EST 12-JAN-2004
DEFINITION	892976 MRC 4Pig Sus scrofa cDNA 5' mRNA sequence.

ACCESSION	CK449409	
VERSION	CK449409.1	GI:40796623
KEYWORDS	EST.	
SOURCE	Sus scrofa (pig)	
ORGANISM	Sus scrofa	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.	
REFERENCE	1 (bases 1 to 940)	
AUTHORS	Smith, T.P.L., Freking, B.A., Ford, J.J., Vallet, J.L., Wise, T.A.,	
	Nonneman, D.S., Wray, J.E. and Keele, J.W.	
TITLE	Porcine EST collection using a normalized library constructed from	
	embryos representing early developmental stages	

ORIGIN

Query Match 27.4%; Score 745.2; DB 7; Length 940;
Best Local Similarity 86.6%; Pred. No. 7.6e-205;
Matches 876; Conservative 0; Mismatches 63; Indels 73; Gaps 2;

RESULT	9
BGI96708	
LOCUS	BGI96708 800 bp mRNA linear EST 21-APR-2001
DEFINITION	RST15934 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.


```
ACCESSION      BG196708
VERSION        BG196708.1  GI:13718395
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 800)
AUTHORS        Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
               Cain,S., Leventhal,C., Thornton,M., Ramachandran,P.K.,
               Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
               Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
               Offenbacher,J., Danzig,J. and Ducar,M.
TITLE          Creation of genome-wide protein expression libraries using random
               activation of gene expression
JOURNAL        Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE        21227151
PUBMED        11329013
COMMENT        Contact: Scott J. Cain
               Athersys, Inc. Ave, Cleveland, OH 44115, USA
               3201 Carnegie Ave,
               Tel: 216 431 9900
               Fax: 216 361 9596
               Email: scaine@atersys.com
               High quality sequence stop: 481.
               Location/Qualifiers
                 1..800
                   /organism="Homo sapiens"
                   /mol_type="mRNA"
                   /db_xref="taxon:9606"
                   /cell_line="HT1080"
                   /clone_lib="Athersys RAGE Library"
                   /note="See 'Creation of Genome-wide Protein Expression
                   Libraries using Random Activation of Gene Expression',
                   Nature Biotechnology, in press. Note that even though the
                   cell type indicated is HT1080, since a random activation
                   method was used, these sequence tags are not necessarily
                   expressed in HT1080 under normal circumstances."
ORIGIN
Query Match      27.1%; Score 738.6; DB 4; Length 800;
Best Local Similarity 98.9%; Pred. No. 5.9e-203;
Matches 796; Conservative 0; Mismatches 4; Indels 5; Gaps 5;
Qy 1210 AATTGGCTTGGAAACAAATTCCTATTATTCACCCCAATGCATTTTCCACTTTGCCATCC 1269
Db 1 AATTGGCTTGGAAACAAATTCCTATTATTCACCCCAATGCATTTTCCACTTTGCCATCC 60
Qy 1270 CTAATAAAGCTGGACCTATCGTCCAACTCGTGTCTCTTTTCCCTATAACTGGGTTACAT 1329
Db 61 CTAATAAAGCTGGACCTATCGTCCAACTCGTGTCTCTTTTCCCTATAACTGGGTTACAT 120
Qy 1330 GGTTTAACTCACTTAAATAAATTAACAGGAATCATGCTTTACAGAGCTTGATATCATCTGAA 1389
Db 121 GGTTTAACTCACTTAAATAAATTAACAGGAATCATGCTTTACAGAGCTTGATATCATCTGAA 180
Qy 1390 AACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGCTGTGCATTTCGA 1449
Db 181 AACTTTCCAGAACTCAAGGTTATAGAAATGCTTATGCTTACCAGTGCTGTGCATTTCGA 240
Qy 1450 GTGTGTGAGATGCCTATAAGATTTCTAATCAATGGAATAAGGTGACAAACAGCAGTATG 1509
Db 241 GTGTGTGAGATGCCTATAAGATTTCTAATCAATGGAATAAGGTGACAAACAGCAGTATG 300
Qy 1510 GACGACCTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGTCGACTTGAA 1569
Db 301 GACGACCTTCATAAGAAAGATGCTGGAATGTTTTCAGGCTCAAGATGAACGTCGACTTGAA 360
Qy 1570 GATTTCCTGCTTGACTTTGAGAGACCTGAAGCCCTTCATTCAGTCAGTGTTTCACCT 1629
Db 361 GATTTCCTGCTTGACTTTGAGAGACCTGAAGCCCTTCATTCAGTCAGTGTTTCACCT 420
Qy 1630 TCCCCAGGCCCTTCAAACCCCTGTGAACACCTGCTTGATGGCTGCTGATCAGAATTGGA 1689
```

```
Db 421 TCCCCAGGCCCTTCAAACCCCTGTGAACACCTGCTTGATGGCTGCTGATCAGAATTGGA 480
Qy 1690 GTGTGGACCATAGCAGTTCTGGCAGCTTACTTGTATGCTTTGGTACCTTCAACAGTTTC 1749
Db 481 GTGTGGACCATAGCAGTTCTGGCAGCTTACTTGTATGCTTTGGTACCTTCAACAGTTTC 540
Qy 1750 AGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTTCATCGCAGCAGTGAAC 1809
Db 541 AGATCCCTCTGTACATTTCCCCCAATTAACCTGTTAAATGGGGTTCATCGCAGCAGTGAAC 600
Qy 1810 ATGCTCAGCGAGAGTCTCCAGTCCGCTGCTGCTGGTGGTGGATGCTTCACTTTGGCAGC 1869
Db 601 ATGCTCAGCGAGAGTCTCCAGTCCGCTGCTGCTGGTGGTGGATGCTTCACTTTGGCAGC 660
Qy 1870 TTTCAGACATGCTGCTGGTGGAGAAATGGGGTGGTTCGCCATGTCATTTGTTTGTG 1929
Db 661 TTTCAGACATGCTGCTGGTGGAGAAATGGGGTGGTTCGCCATGTCATTTGTTTGTG 720
Qy 1930 TCCATTTTTCCTCAGAAATCATCTGTTTTCCTGCTTACTCTGGCAGCCCTGGAGCGTGG 1989
Db 721 -CCATTTTTCCTCAG-ATCATCTGTTTTCCTG-TTACTCTGGC-GGCTCGAGCGTGG 776
Qy 1990 TTCTCTGTGAATATTTCTGCAAAAT 2014
Db 777 TTCTCTGTG-AATATTCTGCAAAAT 800

RESULT 10
LOCUS    BM795259
DEFINITION K-EST0077025 S21SNU520 Homo sapiens cDNA clone S21SNU520-31-C01 5',
           mRNA sequence.
ACCESSION BM795259
VERSION    BM795259.1  GI:19143491
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 671)
AUTHORS    Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
           Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
           Kim,Y.S.
TITLE      21C Frontier Korean EST Project 2001
JOURNAL    Unpublished (2002)
COMMENT    Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: yongsung@mail.kribb.re.kr
           Plate: 31 row: C column: 01
           High quality sequence stop: 671.
           Location/Qualifiers
             1..671
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="S21SNU520-31-C01"
               /sex="F"
               /tissue_type="Stomach"
               /cell_type="Floating aggregates"
               /cell_line="SNU-520"
               /lab_host="Trop10F"
               /note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
               Site 2: NotI; The poly (A)+ RNA was dephosphorylated with
               bacterial alkaline phosphatase (BAP) and then decapped
               with tobacco acid pyrophosphatase (TAP). The decapped
               intact mRNA was ligated with DNA-RNA linker including EcoR
               I site by treatment of T4 RNA ligase and the first strand
```


cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transposition of competent cells E. coli Top10F⁺ by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library."

ORIGIN

Query Match 24.6%; Score 669.4; DB 4; Length 671;
Best Local Similarity 99.9%; Pred. No. 8e-183;
Matches 670; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1775 TTAACCTGTTAATTGGGTCATCGCAGCATGCAATGCTCACGGAGTCTCCAGTGGCG 1834
Db 1 TTAACCTGTTAATTGGGTCATCGCAGCATGCAATGCTCACGGAGTCTCCAGTGGCG 60

QY 1835 TGCTGGCTGGTGGATGCGTTCACCTTTTGGCAGCTTTTGCCAGCATGCTGCTGGTGGG 1894
Db 61 TGCTGGCTGGTGGATGCGTTCACCTTTTGGCAGCTTTTGCCAGCATGCTGCTGGTGGG 120

QY 1895 AGAATGGGGTGGTGGCCATGATGTTGGTTCATTTTGGTTCAGTATCATCTG 1954
Db 121 AGAATGGGGTGGTGGCCATGATGTTGGTTCATTTTGGTTCAGTATCATCTG 180

QY 1955 TTTTCTCTGTTACTTGGCAGCCCTGGAGCGTGGTCTCTGTGAATATTTGCAAAAT 2014
Db 181 TTTTCTCTGTTACTTGGCAGCCCTGGAGCGTGGTCTCTGTGAATATTTGCAAAAT 240

QY 2015 TTGAACGAAAGCTCCATTTTCTAGCCTGAAAGTAAATCATTTTGTCTGCTGCTGGTGG 2074
Db 241 TTGAACGAAAGCTCCATTTTCTAGCCTGAAAGTAAATCATTTTGTCTGCTGCTGGTGG 300

QY 2075 CTTGACCATGCGCGAGTTCCCTGCTGGGTGGCAGCAAGTATGCGCGCTCCCTCTCT 2134
Db 301 CTTGACCATGCGCGAGTTCCCTGCTGGGTGGCAGCAAGTATGCGCGCTCCCTCTCT 360

QY 2135 GCCTGCTTGGCTTTTGGGAGCCGACCATGGCTTACATGCTGCTCTCATCTTGC 2194
Db 361 GCCTGCTTGGCTTTTGGGAGCCGACCATGGCTTACATGCTGCTCTCATCTTGC 420

QY 2195 TCAATTCCTTGGCTTCTCATGATGACCATGTCCTTACACCAAGCTCTACTGCAATTTGG 2254
Db 421 TCAATTCCTTGGCTTCTCATGATGACCATGTCCTTACACCAAGCTCTACTGCAATTTGG 480

QY 2255 ACAAGGAGACCTGGAGAAATATTTGGGACTGCTCTATGTAATAACACATTTGCGTGGC 2314
Db 481 ACAAGGAGACCTGGAGAAATATTTGGGACTGCTCTATGTAATAACACATTTGCGTGGC 540

QY 2315 TCTTCAACACATGCTCTTAACCTGCTGCTGCTTCTTGTGCTTCTCTCTCTTTAATAA 2374
Db 541 TCTTCAACACATGCTCTTAACCTGCTGCTGCTTCTTGTGCTTCTCTCTCTTTAATAA 600

QY 2375 ACCTTACATTTTACGTCTGAAGTAATTAAGTTTATCTTCTGCTGGTAGTCCCACTTC 2434
Db 601 ACCTTACATTTTACGTCTGAAGTAATTAAGTTTATCTTCTGCTGGTAGTCCCACTTC 660

QY 2435 CTGCAATGCTC 2445
Db 661 CTGCAATGCTC 671

RESULT 11

BC038795
LOCUS
DEFINITION
Homo sapiens leucine-rich repeat-containing G protein-coupled receptor 6, mRNA (cDNA clone IMAGE:5220507), with apparent retained intron.
ACCESSION
BC038795
VERSION
BC038795.1 GI:24433474

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

HTC.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 3509)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

HTC.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 3509)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

HTC.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 3509)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

HTC.

Homo sapiens (human)

Homo sapiens

Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;

1 (bases 1 to 3509)

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

1138	-----GACCTAAGACATATGAATAATCTACGAAATTTAAAGTTTGAC	1178
1107	AGCCTCTTGTGCTAGCGCCTCCAAACAACCGCATCTGGGAAATTTGGAGCTGAC	1166
1177	ACTTTCCAGCAGTTGCTTAGCCTCCGATCGCTGAAATTTGGCTTGGACAAAATTTGCTATT	1236
1167	ACCTTCAGCAGCTGAGCTCCCTGCAAGCCCTGGATCTTTAGCTTGGAAACGCATCCCGGTCC	1226
1237	ATTACACCCAAATGCAATTTTCCACTTTGGCCATCCCTAATAAAGCTTGACACCTATCGTCCAAC	1296
1227	ATCCACCTTGAGGCTTTCTCCACCTGCACTCCCTGGTCAAGCTGCACTGACAGACAAC	1286
1297	CTCCTGCTGCTTTTCTTAACTATGAGTTTACATGTTTAACTCACTTTAAATTTAAACAGGA	1356
1287	CAGCTGACCACACTGCCCCCTGGCTGGACTTGGGGCTTGATGCTGAAGCTCTCAAGGG	1346
1357	AATCATGCTTACAGAGCTTGATATCATCTGAAAACCTTTCCAGAACTCAAGTTATAGAA	1416
1347	AACTTTGCTCTCTCCCAAGGCTTTCTCCAAAGGACAGTTTCCCAAACTGAGGATCCCTGGAG	1406
1417	ATGCCCTTATGCTTACCAGTGTGTGCAATTTGGAGTGTGTGAAGATGCCCTATAGAATTTCT	1476
1407	GTGCCCTTATGCCCTACAGTGTGTCTCCATAGGATGTGTGCCAGCTTCTTCAAGGCTCT	1466
1477	AATCAATGGAATAAAGGTGACAAACAGCAGTATGGACGACCTTTCATAGAAGAAGAA	1529
1467	GGGACGTGGGAGGCTGAAGACCTTCACTTGTATGATGAGGAGTCTTCAAAAAGGCCCTCG	1526
1530	-----TGCTGGAATGTTTCAAGGCTCAAGATGAACGTGACCTTGAAGATTTCTGCTT	1581
1527	GGCTCTCTTGGCCAGACAAGCAGAGAAACCATATGACACGAGACCTTGGATGAGCTTCCAGCTG	1586
1582	GACTTTGAGGAGACCTGAAAGCCCTTCATTCACTGACAGTGTTCACCTTCCCAGAGCCCC	1641
1587	GAGATGGAG---GACTCAAAAGCCACACCCAGTGTCCAGTGTAGCCCTACTCCAGGCCCC	1643
1642	TTCAAAACCTGTGAAACACCTGTGATGGCTGCTGATCAGAAATTTGAGTGTGGACCATTA	1701
1644	TTCAAGCCCTGTGAGTACCTTTTGAAGCTGGGGCATCCGCTGCGCGTGTGGGCGCATC	1703
1702	GCAGTTCTGGCACTTACTTGTAAATGCTTTTGGTGACTTCAACAGTTTTTC---AGATCCCCT	1758
1704	GTGTTGCTCTCCGTGCTGCAATGGACTGGTCTGCTGACCGTGTTCGCTGGCGGGCCT	1763
1759	CTGTACATTTCCCCCAATAAACCTGTTAAATTTGGGGTTCATCGCAGCAGTGAACATGCTCAGC	1818
1764	GGCCCCCTGCCCCCGGTCAAGTTTGTGATGATGCGATTTGCAGGCGCCACACCTTGACT	1823
1819	GGAGTCTCCAGTGCCTGCTGGCTGGTGTGGATGCGTTTCACTTTTGGCAGCTTTGCACGA	1878
1824	GGCATTTCTGTGGCCTCTTAGCCTCAGTCGATGCCCTGACCTTTGCTCAGTTCTCTGAG	1883
1879	CATGGTGCCTGTGGGAGAAATGGGGTTGGTGGCATGTCAATGGTTTTTTTTTGTCATTTTT	1938
1884	TACGGAGCCCGCTGGGAGACGGGGCTAGGCTGCGGGGCCATGCGCTTCTCTGGCAGTACTT	1943
1939	GCCTTCAGAAATCACTCTTTTTCTGCTTACTCTGGCAGCCCTCGAGCGTGGTTCTCTGTG	1998
1944	GGGTGCGAGGCACTCGGTGCTGCTCACTCTGGCCGAGTGCAGTGTGACGCTCTCGTC	2003
1999	AAATAATCTTGCAAAATTTGAAAACGAAAGCTTCAATTTTCTAGCCTGAAAGTAATCATTTTTG	2058
2004	TCCTGTGTCGGGCCCTATGGGAAGTCCCCCTCCCTGGGAGCGTTTCAGCAGGGGTCCCTA	2063
2059	CTCTGTGCCCTGTGCGCTTGACCATGTGGCGCAGTTTCCCCCTGTGGGTGGCAGCAAGTAT	2118
2064	GGCTGCTGCGCACTGCGAGGGCTGGCGCGCGCACTGCCCTCGCTCAGTGGGAGAAATAC	2123
2119	GGGCGCTCCCCTCTCTGCGCTGCGCTT-----TGCCCTTTGGGAGAGCCAGCACTATGGGC	2172
2124	GGGGCTTCCCCCACTCTGCGCTGCCCTTAGCGGCCACCTTAGAGGTTCAGCCAGCAGCCCTTGGC	2183

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 873)
AUTHORS Li, W.B., Gruber, C., Jesses, J. and Polayes, D.
TITLE Full-length cDNA libraries and polymorphisms
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced qi:30619059.

Genoscope - Centre National de Sequencage
Concact: Genoscope
2 rue Gaston Cremieux, CP 5708 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
10101.f
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?ss=CS2AAW3ZE04> AW9ZAO3 1sc=10101.f.

```

FEATURES
source
1. 873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D1029YG09"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

```

ORIGIN

Query Match	21.5%	Score 585.4	DB 5	Length 873
Best Local Similarity	95.3%	Pred. No. 2.6e-158		
Matches 646	Conservative 0	Mismatches 27	Indels 5	Gaps 4
Qy	1	ATGAGACACCTCCCGGCTCGGTGTGCTCCTGTGCTTGGCTGTGCTGTGCTGACGTGGCGACC	60	
Db	138	ATGAGACACCTCCCGGCTCGGTGTGCTCCTGTGCTTGGCTGTGCTGTGCTGACGTGGCGACC	197	
Qy	61	GGGGGACAGCTCTCCACAGGTCTGGTGTGTGTGTCAGAGGGGCTGCCCCACACACTGTCAATTGC	120	
Db	198	GGGGGACAGCTCTCCACAGGTCTGGTGTGTGTGTCAGAGGGGCTGCCCCACACACTGTCAATTGC	257	
Qy	121	GAGCCCGACGCGAGGATGTGCTCAGGGTGGACGTGCTCCGACTGGGGCTCTCGAGCTG	180	
Db	258	GAGCCCGACGCGAGGATGTGCTCAGGGTGGACGTGCTCCGACTGGGGCTCTCGAGCTG	317	
Qy	181	CCTTCCAACTCAGCGCTTCACCTCTCACTAGACCTCAGTATGAAACAACATCAGTCAG	240	
Db	318	CCTTCCAACTCAGCGCTTCACCTCTCACTAGACCTCAGTATGAAACAACATCAGTCAG	377	
Qy	241	CTGCTCCCGAATCCCTCGCCAGTCTCCGGTCTTCGTGAGGAGGTTCACGCTTTGGGGGAAAC	300	
Db	378	CTGCTCCCGAATCCCTCGCCAGTCTCCGGTCTTCGTGAGGAGGTTCACGCTTTGGGGGAAAC	437	
Qy	301	GCTCTGACATACATTCGCAAGGAGCAATCACTGGCCCTTTACAGCTCTTAAAGTCTCTTATG	360	
Db	438	GCTCTGACATACATTCGCAAGGAGCAATCACTGGCCCTTTACAGCTCTTAAAGTCTCTTATG	497	
Qy	361	CTGCAGAAATAACAGCTTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT	420	
Db	498	CTGCAGAAATAACAGCTTAAGACAGTACCCACAGAGCTCTGCAGAAATTTGCGAAGCCTT	557	
Qy	421	CAATCCCTCGGTCTGGATGCTTAAACACATCAGCTATGTGCCCCCAAGCTGTTCAGTGGC	480	
Db	558	CAATCCCTCGGTCTGGATGCTTAAACACATCAGCTATGTGCCCCCAAGCTGTTCAGTGGC	617	
Qy	481	CTGCATTTCCCTGAGGCACTGTGTGCTGGATGACAAATGCGTTAAACAG-AAATCCCCCGTCCA	539	
Db	618	CTGCATTTCCCTGAGGCACTGTGTGCTGGATGACAAATGCGTTAAACAGAAATCCCCCGTCCA	677	

Qy	540	GGCTTTTGAAGTTTATCGG-CATTGCAAGCCATGACCTTGGCCCT-GAACAAAATACAC	597
Db	678	GGCTNTTTGAAGTTTATTTGGCCCATGCAAGCCATGACCTTGGCCCTGGTACATAATACAC	737
Qy	598	CACATACCAGACTATGCTTTTGGAAACCTCTCCAGCTT--GGTAGTTTCTACATCTCCATA	655
Db	738	CACATCCAGACTTTGGCTTTGGAAACCTCTCCAGCTTGGCAGGCTTCCATCTCCCTC	797
Qy	656	ACAATAGATCCACTCC 673	
Db	798	ACATTAGGAATCCCACTC 815	
RESULT 14			
AK044357			
LOCUS			
DEFINITION			
ACCSSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
PUBMED			

1216 -----TTATGCAAAATTTAAACACAGAGATAACAGCCCCAGACCAACAGATGTG 1264
QY AAGAAAGATGCTGGATGTTTCAGGCTCAAGATGACGTGACCTTGAAGATTTCTGCTT 1581
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 1265
ACAAAGA-----GAAAGGTGCTACAGATGACGAAATGCCACCAGC 1306
QY GACTTTGAGGAAGACCTGAAAGCCCTTCATTTCAAGTGACGTGTTCACTTCCAGGCCCC 1641
Db ACTGCTGAAGTGAAGACATAGCAAAATATCATCCATTTGACACCTTCAACAGGTGCT 1366
QY TTCAAAACCTGTGAACACCTGCTTGAAGGCTGGCTGATCAGAAATGGAGTGTGGACCATTA 1701
Db TTTAAGCCCTGTGAATATTTACTGGGAAGCTGGATGATTCGCCCTTACAGTGTGGTTCAAT 1426
QY GCAGTTCTGGCACTTACTTGTAAAGCTTTGGTGAATTCACAGTTTTCAGATCCCTCTG 1761
Db TTTCTGGTGGCTTCTTTCACCTGCTTGTCAATTTTAAACAGTGTTCGCTCTGTTTCA 1486
QY TACATTTCCCCCATTTAAACTGTTAATTTGGGTGTCATCGCAGCAGTGAACATGCTCACGGGA 1821
Db TCACTGCTGCTCCAGCTCTTCATAGCTTGTATTCGTGTCTAATCTACTCATGGGC 1546
QY GTCTCCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1881
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| 1606
ATCTATACCTGCTCTTACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1606
QY GGTGCTGCTGGGAGAAATGGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1941
Db GGCATTTGGTGGGAAACTGGCAGCGGCTGCAAGGTAGCTGGGTCTCTGGCAGTCTTCTCC 1666
QY TCAGAAATCATCTGTTTCTGCTTACTCTGCGAGCCCTGGAGCGTGGTCTCTGTGAAA 2001
Db TCAGAGAGCGCTGATTTCTGTTAACCTGGAGCGCTGGAAAGAGCGTTTTTGGCAAG 1726
QY TATTCGCAAAATTTGAAACGAAAGCTCCATTTTCTAGCCTGAAAGTAAATCATTTTGTCT 2061
Db GATGTAATGAAATTTGGGAAAGCAGTCACCTCGACAGTTCCAGTGGCTGCTCGTA 1786
QY TGTGCTGCTGGCTTGAACATGCGCGAGTTCCTGCTGGTGGGAGCAAGTATGGC 2121
Db GCTTTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1846
QY GCCTCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2181
Db GCATCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1906
QY GCTCTCATCTTGTCTCAATTTCCCTTTGCTTCTCATGATGACCATTTGCCCTACCAAGCTC 2241
Db ACCCTAGTGTATTAACCTACTAGCATTTTATTTGATGGCCATTTATCTACACTAAACTC 1966
QY TACTGCAATTTGGACAGGGAGCTGGAGATATTTGGGACTGCTCTATGGTAAACAC 2301
Db TACTTCAACTTAGAGAAAGAGACCCGTGAGAAACTCCCACTGATGATTAAGCAC 2026
QY ATTGCTCTGTTGCTCTTCAACCACTGCTATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2361
Db GTTGTGCTGCTCATCTTCAAAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2086
QY TCCTCTTTTAAACCTTTTACATTTATCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2421
Db GCACCATTTGATCAGCGCAATCTCCATCATGCGCCCGAGATAAATGAAGTCTGTTACGCTGATA 2146
QY GTAGTCCCACTTCTGCAATGCTCAATCCCTTCTCTACATCTTGTTCATTCCTCACTT 2481
Db TTTCTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2147
QY AAGGAGGATCTGGTGGCTGTA 2503
Db AAAGACGACTGGAAGCTCTGTA 2228

LOCUS CD617354 590 bp mRNA linear EST 12-JAN-2004
DEFINITION 55041415H1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION CD617354
VERSION CD617354.1 GI:40265619
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (Bases 1 to 590)
AUTHORS Fu, G.K., Wang, J.T., Yang, J., Au-Young, J. and Stuve, L.L.
TITLE Circular rapid amplification of cDNA ends for high-throughput
extension cloning of partial genes
JOURNAL Genomics 84 (1), 205-210 (2004)
COMMENT Contact: Pu GK
Incyte Genomics, Inc.
3160 Porter Dr., Palo Alto, CA 94304, USA
Tel: 6508454102
Email: gfu@incyte.com.
FEATURES
source Location/Qualifiers
1..590
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="FLP"
/note="Vector: pDrive Cloning Vector"
ORIGIN
Query Match 21.0%; Score 573.4; DB 6; Length 590;
Best Local Similarity 99.8%; Pred. No. 7e-155;
Matches 574; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 292 GCGGGAACGCTCGACATACATTTCCCAAGGGAGCATTTCACTGGCCTTACAGTCTTAAA 351
Db 16 GCGGGAACGCTCGACATACATTTCCCAAGGGAGCATTTCACTGGCCTTACAGTCTTAAA 75
QY 352 GTTCTTATGCTGCAGAAATAATCAGCTAAGACACAGTACCACAGAAAGCTCTGCAGAAATTG 411
Db 76 GTTCTTATGCTGCAGAAATAATCAGCTAAGACACAGTACCACAGAAAGCTCTGCAGAAATTG 135
QY 412 CGAAGCTTCAATCCCTGGCTCGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGT 471
Db 136 CGAAGCTTCAATCCCTGGCTCGATGCTAAACACATCAGCTATGTGCCCCCAAGCTGT 195
QY 472 TTCACTGGCTGCAATCCCTGAGGACCTGCTGGATGACATGCTTAAACAGAAATC 531
Db 196 TTCACTGGCTGCAATCCCTGAGGACCTGCTGGATGACATGCTTAAACAGAAATC 255
QY 532 CCCGTCCAGGCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAACAAA 591
Db 256 CCCGTCCAGGCTTTTAGAAGTTTATCGGCATTTGCAAGCCATGACCTTGGCCCTGAACAAA 315
QY 592 ATACACACATACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTC 651
Db 316 ATACACACATACAGACTATGCTTTGGAAACCTCTCAGCTTGGTAGTTCTACATCTC 375
QY 652 CATACCAATAGATCCCTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCCTACAG 711
Db 376 CATACCAATAGATCCCTCCCTGGGAAAGAAATGCTTTGATGGCTCCACAGCCTACAG 435
QY 712 ACTTTAGATTTAAATTTACAATAAATTTGATGAATTTCCCACTGCAATTTAGGACACTCTCC 771
Db 436 ACTTTAGATTTAAATTTACAATAAATTTGATGAATTTCCCACTGCAATTTAGGACACTCTCC 495
QY 772 AACCTTAAAGAACTAGGATTTTATAGCAAAATATCAGGTGCTGATACCTTGAGAAAGCATTT 831
Db 496 AACCTTAAAGAACTAGGATTTTATAGCAAAATATCAGGTGCTGATACCTTGAGAAAGCATCT 555
QY 832 GTAGGCAACCTCTCTCTTATTACAAATACATTTCTA 866
Db 556 GTAGGCAACCTCTCTCTTATTACAAATACATTTCTA 590

Search completed: July 12, 2005, 20:26:50
Job time : 9772.8 secs

This Page Blank (uspto)